Mathematical models for sustainable wealth redistribution

JC van der Walt

Thesis presented in partial fulfilment of the requirements for the degree of
Master of (Industrial) Engineering
in the Faculty of Engineering at Stellenbosch University

Supervisor: Prof JH van Vuuren

December 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 1, 2017

Abstract

Economic inequality has increased in most large free-market economies during the last century and it has been suggested that this phenomenon is an inherent feature of free-market activities. It seems self-evident, however, that a continual rise in economic inequality is unsustainable. In fact, severe economic inequality has historically been associated with negative effects such as poor economic growth, severe financial recessions or even violent revolutions. Wealth redistribution is present in every form of government, although the extent thereof varies, and existing theoretical justifications of redistributive actions usually rely heavily on utility theory.

Most economic postulates related to inequality are empirically inspired and defended, but because of the vast variety of possible economic contexts in which they may prevail, many of these claims are disputed. One example is the so-called *Robin Hood paradox*, which asserts that the extent of wealth redistribution is less in more unequal societies, where it is needed most, than in more economically equal societies. Another is the *Kuznets curve*, which predicts that the extent of inequality in a developing economy will follow an inverted 'u' curve as a result of development over time.

The implications of increasing relative inequality over time as an inherent feature of wealth growth are investigated in the presence of wealth redistribution. Very simple mathematical model abstractions are employed to shed light on the possible evolution over time of wealth distribution in the context of very basic assumptions, since such behaviour may perhaps then also be inferred in more complicated settings.

Assuming increasing per capita wealth growth-rate functions is one way of capturing increasing relative inequality over time, the very simplest case being linearly increasing per capita wealth growth-rate functions, which are considered in this thesis. Two examples of redistribution dynamics are investigated. One example mimics diffusion-like effects of trickle-down redistribution, while the other represents a conservative, linear-tax transfer scheme.

It is established analytically within the context of the aforementioned mathematical models that increases in economic inequality can always be limited by means of sufficient redistribution. It is also demonstrated that the Robin Hood paradox may follow from very simple assumptions. It is furthermore illustrated that fluctuating behaviour in the evolution over time of wealth inequality can even manifest itself in the absence of time-dependent processes, and hence that explanations of such trends which merely assume time-dependent underlying processes might be of dubious value. Examples of analytical formulations of theoretical justifications for redistributive actions, independent of utility theory, are finally also provided.

Uittreksel

Ekonomiese ongelykheid het gedurende die laaste eeu in die meeste groot vryemark ekonomieë toegeneem en daar is al voorgestel dat hierdie verskynsel 'n inherente eienskap van vryemark aktiwiteite is. Dit blyk egter voor die hand liggend te wees dat 'n voortdurende toename in ekonomiese ongelykheid onvolhoubaar is. Ernstige ekonomiese ongelykheid het trouens histories hand aan hand gegaan met newe-effekte soos stadige ekonomiese groei, ernstige finansiële resessies en selfs gewelddadige revolusies. Die herverdeling van rykdom is in elke vorm van regering teenwoordig, alhoewel die mate daarvan varieer, en bestaande teoretiese regverdigings vir herverdelingsaksies berus gewoonlik swaar op nutsteorie.

Die meeste ekonomiese postulate wat te make het met ongelykheid is empiries-geïnspireer en word ook sodanig verdedig, maar baie van hierdie bewerings word as gevolg van die groot verskeidenheid moontlike ekonomiese kontekste waarin hul mag voorkom, betwis. Een voorbeeld hiervan is die sogenaamde *Robin Hood-paradoks* waarvolgens die mate van rykdom-herverdeling minder is in meer ongelyke gemeenskappe, waar dit juis méér benodig word, as in ekonomies meer gelyke gemeenskappe. Nog 'n voorbeeld is *Kuznets se kromme* waarvolgens voorspel word dat die mate van ongelykheid in 'n ontwikkelende ekonomie 'n omgekeerde 'u' kromme sal volg soos ontwikkeling oor tyd geskied.

Die gevolge van toenemende relatiewe ongelykheid oor tyd as 'n inherente kenmerk van toenemende rykdom, in die teenwoordigheid van rykdom-herverdeling, word in hierdie tesis ondersoek. Baie eenvoudige wiskundige modelabstraksies word ingespan om lig te werp op die moontlike evolusie oor tyd van die verdeling van rykdom in die konteks van baie basiese aannames, aangesien sodanige gedrag dan ook moontlik in meer ingewikkelde kontekste afleibaar is.

Die aanname van toenemende per kapita groeitempo-funksies is een manier waarop die toename in relatiewe ongelykheid oor tyd vasgevang kan word. Die heel eenvoudigste geval hiervan is lineêr-toenemende per kapita groeitempo-funksies, wat in hierdie tesis oorweeg word. Twee voorbeelde van herverdelingsdinamika word ondersoek. Een voorbeeld boots die diffusieverwante gedrag van deursyferingsherverdeling na, terwyl die ander 'n konserwatiewe, lineêre belasting-oordragskema is.

Daar word binne die konteks van die bogenoemde wiskundige modelle analities vasgestel dat toenames in ekonomiese ongelykheid altyd deur genoegsame herverdeling beperk kan word. Daar word ook aangetoon dat die Robin Hood-paradoks die gevolg van baie eenvoudige aannames mag wees. Verder word daar gedemonstreer dat wisselende gedrag in die evolusie van rykdom-ongelykheid oor tyd selfs in die afwesigheid van tyd-afhanklike prosesse mag voorkom en gevolglik dat verklarings van sulke tendense waarin onderliggende tyd-afhanklike prosesse bloot aangeneem word, van twyfelagtige waarde mag wees. Voorbeelde van analitiese formule-rings vir die teoretiese regverdiging van herverdelingsaksies word laastens ook buite die konteks van nutsteorie gegee.

Acknowledgements

The author wishes to acknowledge the following people and institutions for their various contributions towards the completion of this work:

- Professor van Vuuren for his unwavering support, meticulous feedback and friendship.
- The National Research Foundation for funding this research.
- The SUnORE research group for the provision of office space, computing facilities and thoroughly enjoyable social events.

Table of Contents

List of Reserved Symbols

Symbols in this thesis conform to the following font conventions:

$\boldsymbol{a}$	Symbol denoting a **vector**	(Italic lower-case letters in boldface)
$\boldsymbol{A}$	Symbol denoting a **matrix**	(Italic capitals in boldface)
$\mathcal{A}$	Symbol denoting a **set**	(Calligraphic capitals)

Symbol	Meaning
μ	The mean value of a real-valued function on an interval
V	The variance of a real-valued function on an interval
$\overline{V}$	The normalised (by the mean) variance of a real-valued function on an interval
$w(\lambda, t)$	A function representing wealth at a position λ and time t in some economic system
$\mathbb{R}$	The set of real numbers
$\mathbb{R}^m$	The set of real-valued m-component (column) vectors
$\mathbb{R}^m_+$	The set of positive real-valued m-component (column) vectors

List of Figures

List of Tables

CHAPTER 1

Introduction

Contents

1.1 Background

Rising economic inequality seems to be an inherent feature of large free markets [117]. Indeed, economic inequality has increased over the last 150 years in both the United States of America and the United Kingdom [91]. Extreme rises in economic inequality have also been associated with the transition from socialist to free market systems [68, 150]. Yet, historical periods of extreme inequality have been characterised by catastrophic events such as (violent) revolution [94] or severe financial recession [77]. Furthermore, high levels of economic inequality are associated with lower economic growth [8, 51, 108, 112]. This suggests that the current drivers of wealth distribution in large free markets are not sustainable.

Societal inequality and the most beneficial structure of social arrangement have been deliberated for thousands of years [41, 126, 129]. The inherent power instability in social systems is not limited to financial matters. Unchallenged political power, which is certainly related to economic power, also tends to become more concentrated over time [66]. Such concentration in large adaptive systems increases fragility and the risk of large shocks [26, 141], which is one reason why the democratic process is of value as a deterrent of such concentration.

Empirical investigations related to economic inequality and its temporal evolution have resulted in controversial hypotheses pertaining to its dependence on development [86], its relation to redistribution [79, 83, 90] and different means of regulation [116]. A definitive measure of inequality, or a means of ranking different distributions according to their relative inequality, is required if notions of decreased or increased inequality are considered. Popular metrics of economic inequality have, however, been shown to rank the same set of distributions differently [11]. Caution should therefore be exercised when inequality is quantified, by rigorously noting the underlying assumptions of the quantitative metric.

The mathematical modelling of systems emanating from the social sciences poses a very difficult task for a number of reasons, including the complexity of capturing the system entity interactions, the difficulty in identifying all applicable variables as well as the suitable measurement of these variables, and the challenge of isolating natural subsystems. Discussions on and investigations of matters pertaining to societal inequality therefore tend to be philosophically or empirically oriented. The validity of findings in such endeavours is, however, always subject to the context in which the observations are made, and are rarely truly universal. The reason for this phenomenon is that these observations are seldom representative of all possible behaviour, and may furthermore be subject to various biases or measurement errors. It is therefore unsurprising that many of the hypotheses constructed primarily on empirical investigations in economics are disputed. In contrast, a theoretical mathematical investigation of the underlying concepts may provide insights that are universally applicable, albeit in a highly idealised or abstracted fashion.

If the temporal evolution of a wealth or income distribution is assumed to be unstable in the sense that relative inequality increases over time, this implies that wealth increases faster (or decreases slower) in the distribution at higher levels of wealth than at lower levels. Assuming that the per capita[1] growth rates dictating this temporal evolution of wealth are captured by an increasing function of wealth is therefore one sufficient way of modelling increasing relative inequality over time.

If this inherent instability is assumed and it is accepted that the unbounded growth of inequality is unsustainable, then it follows that redistribution of wealth is required to render the situation sustainable. Some redistribution is present in all governance systems [50], although the extent of such redistribution varies substantially from system to system. The manner in which this redistribution transpires, or ought to transpire, is not obvious. Wealth redistribution has been described as a trickle-down process [3, 114], which is mathematically similar to the notion of diffusion. This analogy is both elegant and satisfying since diffusion is a natural equalising force present in many contexts. Over time, for example, a gas distributes itself at equal densities throughout the space available. Another example is the distribution of heat in a homogeneous, isolated material which equalises over time as a result of diffusion. Even the spread over time of a beneficial innovation in society has been likened to diffusion [95]. Another approach to modelling redistribution is to impose a particular functional form on wealth, such as linear tax [124, 131], with negative rates at low wealth levels.

When dynamical systems are modelled in which component densities change over time in the presence of diffusion, based on their distributions over space, so-called *reaction-diffusion equations* result. In such equations, a *reaction term* expresses the growth or decline of a given component, while a *diffusion term* models the movement of entities modelled by the component in the direction of decreasing density. One example of such a model is Fisher's equation [57] for the advance of an advantageous gene in a population, and another is the Newell-Whitehead-Segel equation [105] describing Rayleigh-Bernard convection. A linear type of redistribution may be modelled in a similar manner, by replacing the diffusion term with an appropriate alternative.

The reaction terms in the aforementioned models (and in most one-component reaction-diffusion models) evaluate to zero for some positive density and are negative for all densities larger than this, giving rise to the existence of positive, stable equilibrium states. The reader may be excused for expecting that systems which exhibit positive growth rates for large densities (*i.e.* infinite

[1] The use of the phrase *per capita* here refers to the specific wealth attainment level, and not to specific entities. Hence the product of the function evaluated at a given wealth and that wealth attainment is the average wealth growth rate at that point in the wealth distribution, although this rate is not necessarily achieved by any individual entity at that wealth level.

growth over infinite time) do not exist in practice. There exist several systems that exhibit always positive growth over time intervals of (in foresight) uncertain, and possibly considerable, length. For example, in a well-functioning economy, interest rates (and therefore fiscal growth) are expected to be positive, and the duration of a particular economy's stability may be large. It is, however, important to recognise that component densities in such systems do not grow unbounded in finite time, but merely exceed any pre-specified density threshold when viewed over a long enough period of time. The objection that systems in which growth allowing densities to tend toward infinity do not exist therefore does not apply, since the phenomenon of finite-time blow-up is inherent to this objection.

Numerous systems exist in which competing forces are at work — one tending toward increased concentration and inequality, and the other equalising. In these systems, persistence is often preferred over decrease, and increase over persistence. The distribution of wealth and the concentration of political power are two examples.

1.2 Problem description

The problem considered in this thesis involves the derivation of simple mathematical models of wealth distribution over time based on the notion of instability as a result of increasing per capita growth rates over wealth, and counteracting redistribution. Sufficient conditions for bounded inequality are derived in the context of these models. Possible long-time behaviours of solutions to these models are also investigated and potential implications of these behaviours are interpreted in view of existing postulates in the literature on economic inequality.

1.3 Research objectives

The following seven objectives are pursued in this thesis:

I To *conduct* a thorough survey of the academic literature related to:

- (a) The notions of inequality and redistribution with a particular focus on
  - (i) societal and economic inequality,
  - (ii) postulates related specifically to economic inequality,
  - (iii) the identification of suitable metrics of economic inequality, and
  - (iv) the modelling of wealth distribution over time.
- (b) Mathematical modelling and model solution computation in the context of
  - (i) systems of *partial differential equations* (PDEs) in general, and reaction-diffusion models in particular, and
  - (ii) appropriate numerical solution techniques for initial-boundary value problems involving PDEs.

II To *derive* mathematical models for the temporal dynamics of the distribution of wealth, as affected by growth and redistribution processes (one model being within the realm of reaction-diffusion systems).

III To *establish* suitable metrics of inequality in the context of the models of Objective II.

IV To *present* sufficient conditions for decreasing inequality of wealth in the context of the models of Objective II.

V To *analyse* the temporal evolution of the inequality metrics of Objective III for simple growth functions in the context of the models of Objective II in order to characterise the possible behaviour of solutions with reference to relative inequality.

VI To *interpret* the findings of the analysis carried out in pursuit of Objective V within a socio-economic context and to note their possible implications pertaining to existing postulates in the literature on economic inequality.

VII To *recommend* sensible follow-up work related to the work documented in this thesis which may be pursued in future.

1.4 Scope delimitation

The focus in this thesis is an investigation of the nature of change in wealth inequality over time subject to increasing growth functions of wealth in the presence of redistribution of different magnitudes. The implications of the investigation are interpreted in an economic context in order to demonstrate possible applications of the insights gained, and the observations made should therefore be viewed in this light. It is noted, however, that the implications of the investigation may extend to other domains.

All major ethical theories for social arrangement call for equality of some variable [130]. The word *wealth* is used in this thesis to refer to that variable, and the analysis conducted applies to all spaces where the distribution of that variable is determined by entities' ability to increase it, as well as the extent of redistribution present. Furthermore, an entity's ability to increase its attainment of this variable is assumed to be proportional to its current level of attainment of the variable. On a more abstract level than the monetary sphere, the distribution of power in a community also satisfies these assumptions.

The purpose of the research conducted in this thesis is not to attempt to capture the complexity of actual economies, but rather to elucidate the nature of simple concepts related to the emergence of redistribution phenomena over time in a rigorous manner, albeit in a highly abstracted context.

1.5 Thesis organisation

The second chapter of this thesis is devoted to a brief review of the academic literature related to economic inequality and wealth distribution. First, points of view in the literature on inequality in general and on economic inequality in particular are reviewed briefly, in order to create a context within which economic inequality may be discussed fruitfully. Existing postulates under three related headings, namely trends in economic inequality, the redistribution of wealth, and the notion of optimal redistribution, are then reviewed. Metrics used to quantify the extent of inequality of a wealth distribution are reviewed next in order to establish a context in which the reasons for the particular choices of inequality indices used in this thesis can be elucidated. Existing models of wealth distribution are finally reviewed so as to create a context for the contribution of this thesis.

Certain mathematical prerequisites for following the arguments presented later in this thesis are reviewed in the third chapter. First, a number of basic mathematical definitions are provided, and this is followed by a very brief discussion on the subject of differential equations. A particular class of PDEs, namely reaction-diffusion systems, is then reviewed. A method for analysing the stability properties of equilibrium solutions to certain differential equations is outlined next. Numerical techniques for solution approximation are finally reviewed in the context of initial-boundary conditions involving certain partial differential equations.

In the fourth chapter, two mathematical models for wealth distribution over time are derived. First, the underlying model assumptions are noted and motivated, after which mathematical derivations of the two models are presented. The choice of indices used for quantifying inequality in this thesis is then motivated. The per capita wealth growth-rate functions to be considered in the aforementioned models are put forward next. These per capita wealth growth-rate functions are simple examples of functions which produce increasing relative inequality over time, and capture certain basic economic relationships. Since the aim is to demonstrate that certain interesting solution behaviours are possible in a very simple modelling context, the models, their underlying assumptions and the per capita wealth growth-rate functions, are quite rudimentary. It may be inferred that solution behaviours admitted in such extremely simple contexts are also possible in more complicated settings. The existence and uniqueness of the models' solutions, as well as certain basic solution properties (such as nonnegativity of solutions) are then established. The models' limitations and intended use are finally discussed critically and clarified in a formal model apology.

Sufficient conditions for bounded inequality in solutions to the models of Chapter 4 are established in Chapter 5 by taking the time derivative of the inequality indices and seeking conditions which ensure that these derivatives are negative. Lemmas required in the arguments of this chapter are first established, after which the main results of the chapter are presented. The results are finally illustrated in numerical examples in the penultimate section of the chapter.

The long-time behaviour of solutions to the models of Chapter 4 are considered in Chapter 6. The presentation is structured according to the different per capita wealth growth-rate functions in order of increasing complexity. For a simple linear per capita wealth growth rate, the relation between solution persistence and the extent of redistribution is investigated. The stability of the constant equilibrium solutions is established analytically, and an existence result is established for a nonconstant equilibrium solution to one of the models. The solution behaviour of the same model is then investigated numerically, within the context of a mean-dependent linear per capita wealth growth-rate function. The effect of the redistribution rate on solution stability and on the equilibrium solution size and shape is analysed. Finally, a mean and inequality-dependent linear per capita wealth growth-rate function is considered and possible model solution behaviour is investigated numerically within this context. The insights gained in this chapter are captured in labelled observations, which are considered further in the following chapter.

The seventh chapter is devoted to a discussion on the findings of the analysis of Chapters 4 and 5. First, the question of model validation is considered. The expected and actual model behaviours in the simplest cases are discussed, and one of the aims of the investigation and the type of insights that may be gained from the investigation, are restated. Implications of some of the findings of Chapter 6 for the so-called Robin Hood paradox are then considered. The presence of fluctuations in economic metrics over time, and whether their presence implies the existence of time-dependent driving processes, are then discussed. The notions of an optimal redistribution and a theoretical justification of redistributive actions are finally considered.

The contents of this thesis is summarised in Chapter 8, and this is followed by a critical appraisal of the contribution of this thesis. Possible avenues for future research related to the work documented in this thesis are finally suggested.

Part I

Literature review

CHAPTER 2

Economic inequality

Contents

The literature related to economic inequality is reviewed very briefly in this chapter. First, perspectives are offered on inequality from the point of view of social justice and their applicability to economic inequality is noted. Certain historical trends in economic inequality, as well as subsequent theoretical postulates, are mentioned next. Instruments available for the redistribution of resources are summarised in the section that follows and an empirical observation related to the relationship between economic inequality and the redistribution of wealth is noted. The notion of optimal redistribution is then reviewed, after which the measurement of economic inequality is considered by reviewing certain classes of inequality indices. Existing models of wealth distribution are finally introduced to provide context for the contribution of this thesis.

2.1 Inequality and social justice

The most suitable structure of social arrangement has been a topic of debate for millennia [41, 126, 129]. Part of the definition of such a structure must pertain to the distribution of resources amongst the members of a society, and thus economic inequality is necessarily addressed in these discussions, be it explicitly or implicitly. Perhaps the first well-known work on the subject is Plato's *Republic* [118], which suggests a rigorously state-controlled society, ruled by a so-called philosopher king, as the most desirable form of social organisation. In such a society, the unbounded increase of inequality is explicitly prohibited, since the extent of state control and ownership of resources are complete, and the highest ruling authority is expected to lead a simple life of moderation. This does not mean, however, that men are viewed as equals in all respects. Rather, society is to be structured according to a fixed hierarchy, based on individual inherent endowments. *Politics*, by Plato's pupil Aristotle [132], was the next major contribution to the discourse. In this work, Aristotle initially built on the ideas put forward by Plato, and although he defended the right to ownership of private property, his ideal society is still a rigid one where one's life is determined by position in a hierarchical structure. In both these cases then, the extent of societal inequality, while it may seem extreme, is fixed.

The rigid structures suggested in the aforementioned ideas may remind the reader of the communist regimes of Russia or China [62, 149], and with the benefit of hindsight, seem painfully naive and impractical. It is, however, noteworthy that the earliest thinkers' ideal societies did not allow individuals to amass resources in an unlimited fashion.

There seems to be a shift over the ages from a strict formalistic approach to social organisation, in conjunction with a fixed limited equality, to a more liberal view together with a far reaching equality in human value. Prior to the aforementioned Greek philosophers, Confucian philosophy also advocated the view that not all persons are capable of self-government and that able leaders should control resources [120]. The notion of the universal equality of human worth is thought to have originated with the Christian philosophers, such as Thomas Aquinas [37].

During the age of enlightenment, notable contributions were made by the English philosophers Hobbs and Locke. In his famous *Leviathan*, Hobbs [73] suggested that differences in individual abilities should not necessarily correspond to different benefits in society [37]. The famous French author and philosopher Jean-Jacques Rousseau offered an encompassing treatise of the origins, the evolution and possible consequences of inequality amongst men [126]. He distinguished between natural inequality which results from physical differences between persons, and moral or political inequality which evolves from agreements made between men. He then argued that natural inequality is, as the term suggests, part of the diversities of the human race, and hence not of further philosophical interest. Moral inequality is, however, claimed to be unjustifiable. According to Rousseau, moral inequality emerges in conjunction with the formation of society and the ownership of property, as a result of man's corrupted desire for the domination of others. Ultimately this evolution of moral inequality leads to despotism, the unequal rule of society by a single man with unparalleled wealth. Locke, in contrast, argued that the right to property follows from labour, and in this context he defended the unbridled accumulation of wealth in monetary form [16].

Karl Marx criticised the notion of an equal right to the ownership of property, referring to it as a 'right to inequality' [37]. He also described the capitalist system as fundamentally unstable and ultimately destined to lead to revolution motivated by severe inequality [17]. His suggestions for ensuring acceptable economic equality, however, are considered to place too great a constraint on individual liberty [37].

Perhaps the most noteworthy recent addition to the discourse on social arrangement, from the point of view of social justice, is due to Rawls [122]. Its primary feature is the well-known thought experiment, entitled *the original position*. According to this experiment, the principles of social justice, and therefore the most acceptable structure of a society, is to be decided by persons unaware of their own position in this society. It is argued that, from this point of view, an agreement will be reached which minimises the disadvantages experienced by the least well off. Following this line of argument, two principles of justice are introduced. First, each person should have access to the most extensive basic liberty possible, without infringing upon the liberty of another. Secondly, the inequalities in society should be structured so that the least advantaged derive the greatest benefit from these, and everyone should have equal opportunity to fulfil the mandate of official positions or offices.

The first principle resonates with the ideas of capitalist free markets, while the second may be interpreted to call for regulation and redistribution of resources. Rawls may therefore be described as egalitarian in terms of liberty, but also, in some limited sense, in terms of resources.

For if individual liberty alone is emphasised, inequalities which do not benefit the least well off may (and most likely will) arise.

The argument (and empirically observed phenomenon) that some economic inequality may be desirable since it may be associated with greater economic growth [59], which may ultimately benefit the whole of society, may therefore be in agreement with the theory of Rawls, if it can be shown that the poor ultimately derive the greatest benefit. This can only be the case in a free market where some form of redistribution of resources is present when considering the history of capitalism, and the events that unfold when it is unregulated, or worse, regulated in such a manner that the most well off derive the greatest benefit [113].

Sen [130] emphasised that all ideologies are egalitarian in some sense, the difference being the variable concerned (income, liberty or opportunity to hold office, for example). Moreover, Pikkety [117] has argued that unregulated capitalism is fundamentally unstable, producing ever increasing inequality. If this is accepted, then limited egalitarianism in terms of resource ownership, in other words, forced redistribution of wealth, becomes the only way by which capitalism may be considered sustainable, if the extent of attainable wealth is not explicitly limited. Indeed, as has become apparent, the theme of limited egalitarianism in terms of resource ownership in an ideal society is an old one.

2.2 Trends in economic inequality

Since accurate empirical data of individual wealth positions are typically difficult to obtain, investigations of economic inequality have tended to focus on income inequality [49]. It is, however, expected that a strong relationship exists between income inequality and inequality of wealth, the difference being dictated by aggregate consumption *versus* saving decisions within different population subsets [137]. In fact, it has been observed across several countries that household-wealth inequality is proportional to yet greater than household-income inequality [49].

Kuznets famously postulated that income inequality in a developing economy will follow an inverted 'u' curve [86]. First, income inequality will increase as new developments benefit some limited subset of the population, after which it is expected to flatten as the economy becomes more developed, giving rise to more equally spread opportunities across the population, and ultimately steadily decrease when all compete with equal opportunity of success over time. This may be extended to the expected inequality of wealth in a developing economy, since changes in income inequality can be expected to result in similar changes in the inequality of wealth.

While Kuznets' findings are empirically supported, in the sense that such an inverted 'u'-shaped curve exists in the income-distribution inequality of several countries analysed (specifically the United States of America, the United Kingdom and Germany), critics have argued, and Kuznets himself has conceded [86], that the data considered were limited, and hence that the findings were therefore not conclusive [9]. A myriad of studies investigating this phenomenon followed the initial publication, with most, however, failing to reject the hypothesis [5, 72, 80, 144]. For an example of a data analysis rejecting the Kuznets hypothesis, see [9]. The fact that this claim is still disputed is an example of the limited reach of empirical studies in validating or refuting economic postulates.

During the period of Kuznets' seminal publication, a decrease in income inequality was observed in America, which has become known as the egalitarian revolution. A broader perspective on changes in inequality has since emerged, as improved data became available, and with the benefit of approximately half a century of insights. It has become apparent that inequality has increased over the past 150 years in both the United States of America and the United Kingdom [91]. It has subsequently been suggested that rising inequality is an inherent feature of capitalism, if it is not prevented by state intervention [117]. The extreme rise in wealth inequality experienced in Russia, following the transition from limited personal property under socialism to a market system, is a striking example of this [68, 150].

Pikkety [117] observed that, with the exception of the great wars and major recessions, capital returns tend to be greater than overall economic growth. The result is that wealth continually moves from the working class to the owners of capital, thereby accumulating in a small subset of the population and increasing the inequality of wealth. As a remedy, the imposition of a global wealth tax, in conjunction with very progressive income tax, has been suggested. The main aim of these measures is to prohibit the perpetuation of increasing fortunes in small population subsets, which persists through inheritance.

Empirical investigations on the influence of economic inequality on economic growth claim both positive [59, 88] and negative [36] correlations with economic growth, once again highlighting the non-universality of empirically justified claims in economics. The majority of investigations find a negative correlation between inequality and growth [8, 51, 108, 112]. In support of this trend it has been argued that in democratic societies, high levels of inequality result in political instability [81] and reduced investment [7].

In conventional approaches, on the other hand, it has been argued that economic inequality presents an incentive to work [4], and is therefore related to economic growth. This argument only holds, however, if the extent of inequality and economic conditions are such that a person or business can reasonably be expected to transition to higher wealth positions in the population. With extremely high levels of inequality, a significant rise in wealth position becomes increasingly improbable. It may therefore be concluded that inequality can stimulate economic growth if the level of inequality is below a certain threshold, above which it starts to stifle growth.

The continual rise in economic inequality which has been observed for more than a century furthermore shows that economic growth is often related to increasing inequality. If economic growth results in greater inequality, then the net growth rates achieved at higher wealth levels are greater than those achieved at lower wealth levels. It may therefore be concluded that, on average, per capita growth rates in most economies are increasing functions of wealth.

High levels of economic inequality are undesirable for several reasons. As has been noted, income inequality increases political instability [7] and high levels of inequality have also been associated with lower economic growth [112]. In the absence of redistribution, the resulting rising inequality may lead to catastrophic undesirable events such as (violent) revolution [94] or severe financial recession [77]. Empirical history therefore leads to a conclusion similar to that of the previous section, namely that a certain degree of redistribution is required in order to render liberal capitalism sustainable.

Complete redistribution, even if it were practically possible, is, however, also undesirable for several reasons. First, because growth rates increase over wealth, economic growth may be impacted negatively through redistribution. Secondly, the act of redistribution itself is costly and results in lost wealth, a phenomenon known as *Okun's bucket* [106] (named after the economist Arthur Okun who likened wealth redistribution to transferring water using a leaking bucket).

2.3 The redistribution of wealth

In this section, various instruments that are at a government's disposal for altering the wealth distribution in an economy are reviewed. These instruments include taxation, social services and social transfers [119]. The manner in which these might be applied is finally categorised into two classes, depending on the parts of the wealth distribution considered when determining transfer magnitudes.

Wealth redistribution is present in every democracy, although the extent may vary substantially [50]. The provision of basic public services, such as law enforcement or public health, the primary mandate of the democratic state, is financed primarily through taxation. These services which theoretically contribute equally to the lives of all members of the population, are thus not financed equally by all. This is the case even for a flat tax rate in the presence of wealth inequality, although progressive tax is commonly employed. A very basic type of redistribution is therefore a consequence of the provision of public services [18]. This type of redistribution is generally accepted, as the necessity of such basic services is self-evident. The progressiveness of taxation as a means of funding these services is, however, not generally agreed upon.

Social transfers, through the provision of some minimum income level, or by means of negative income taxes (as suggested by Friedman [60]), or a minimum standard of living, through the provision of non-fiscal social grants, are ways in which wealth may be transferred directly to those at the bottom of the income or wealth distribution. Various forms of taxation are applied in order to extract the wealth required in these transferrals, and for the provision of basic services.

Personal income tax and property tax are typically structured in brackets, with higher earners or assets of greater value being subject to progressively larger percentages of taxation [119]. Wealth tax, as suggested by Pikkety [117], is progressive by definition. Not all forms of taxation are, however, necessarily progressive — corporate taxes for medium sized companies, as well as indirect taxes, are often regressive [119]. It is the combination of the available tax instruments and the characteristics of a particular economy subject to these that ultimately determine whether, and how, redistribution takes place. It has been observed that the net effect of typical combined taxation in developing countries tends to be regressive [65]. This means that the net effect of taxes in these often very unequal societies is an increase in inequality, by being of greater benefit to the wealthy rather than to the poor.

The commonly observed phenomenon that the extent of redistribution present in economically unequal countries is substantially less than the redistribution present in more economically equal countries is known as the *Robin Hood paradox* [90]. Since income and wealth distributions are typically skewed to the left [99, 123] (the median is less than the mean), it is expected that in the presence of a democratic process, more redistribution ought to emerge in more unequal societies. Political economic models with egalitarian political power distributions are therefore biased in the opposite relation than is empirically observed in wealth distributions, since the median voter is the decisive voter in this context [99].

There are two prominent (and at times contradictory) schools of thought related to the development of economies with different redistribution policies which give rise to the *Robin Hood paradox*. These approaches are known as *power resource theory* [83] and *varieties of capitalism* [70]. Each of these attempts to explain the clustering of economies into two main groups. The first is *coordinated market economies*, characterised by strong strong labour unions, large welfare states (and therefore substantial redistribution of wealth) and economic equality. The second is *liberal market economies*, which are characterised by fewer constraints on trade relations, small welfare states and greater levels of economic inequality.

Proponents of *power resource theory* argue that political power is related to capital, and since the concentration of monetary capital is greater than that of labour capital, the owners of monetary capital have greater political influence than the labour class [82]. The disparity in power can be lessened, however, through unionisation. This coordinating ability (or lack thereof) within the labour class, drives the formation of greater welfare states through the extension of social citizenship rights and by limiting the market options available to capitalists. The primary driver of egalitarian processes is therefore the labour class, while capitalists may oppose, consent or at times even support these initiatives, depending on the immediate profit implications.

In contrast, *varieties of capitalism* holds that welfare state formation is influenced primarily by capitalists [79]. In industries where asset-specific skills are required, it is argued that persons are exposed to greater unemployment (or below skill-level employment) risk. In such industries, capitalists drive welfare state formation as a hedge against the negative financial implications of unemployment risks, in order to encourage investment in asset-specific skills. Welfare state formation therefore depends on national skill profiles and the transferability of skills between industries.

The continuation of these differing views highlights the opacity of the driving processes behind economic phenomena. This strengthens the argument for adopting an approach of analysing possible outcomes given the presence of certain dynamics, rather than seeking to uncover the driving forces of complex systems.

In determining exactly how wealth redistribution ought to transpire (for example, how tax brackets, and their corresponding proportional taxation levels, should be structured) one may distinguish between two theoretical approaches to wealth transfer. Wealth redistribution through economic interactions has been modelled in a trickle-down fashion [3, 114], where the local distribution of wealth determines the dynamics of wealth transfers. The entire wealth distribution is therefore not explicitly taken into account when determining local transfers, but since all entities are connected to a certain set of "neighbours" in the wealth distribution, the entire distribution implicitly influences the evolution of wealth transfers over time. By formulating regulations aimed at redistributive results in this context, the trickle-down effect of wealth may therefore be magnified or lessened.

Another approach is to impose a functional form on wealth (not dependent on the wealth distribution itself) which dictates redistributive transfers. An example is where an entity's net contribution or benefit derived from the redistribution scheme depends on its wealth position relative to some population-dependent standard (the average wealth, say). This corresponds to a linear tax function [124, 131] which translates on the income axis according to the average wealth.

2.4 The notion of optimal redistribution

The problem of optimal taxation, where a certain revenue is to be raised by means of taxing income in such a manner that the decrement in utility is minimised, was first formulated analytically in 1927 by the celebrated mathematician F. P. Ramsey [121]. An entire field of economic enquiry, still actively being pursued (see, [34, 125] for example), stemmed from this problem formulation. This field is concerned primarily with raising the revenue required to fund social goods in the most efficient fashion, and not with redistributive transfers which may also form part of a tax system (although some overlap is inevitable, given the interrelation).

The notion of optimal redistribution has been formalised in a similar vein to optimal taxation, also by utilising utility theory. The sovereignty of the individual consumer, together with the notion of independent utility, was central to the emergence of neoclassical public finance [29]. However, redistributive practices observed and evidently required in reality, whether informal charity or formal government intervention, are inconsistent with the aforementioned axioms of neoclassical public finance [74]. The lack of a theoretical justification for redistributive actions within the theory of welfare economics was addressed in the pioneering papers [74, 107] by introducing interdependent individual utility functions. In addition to the notion that a person's utility may be positively related to the welfare of society, security against drastic future income fluctuations were noted as a possible motivation for the redistribution of resources. The latter is, however, not explicitly incorporated in the analytical theory of optimal redistribution. The triumph of this theory is its ability to demonstrate analytically that progressive taxation may be socially optimal, consistent with the Pareto criterion[1], under reasonable conceptual assumptions. It should be noted that interpersonal utility comparisons (which are considered to be impossible [71]) are not required in this argument.

Mirrlees [101] extended the analytical treatment of optimal redistributive income tax, assuming that individual utility is positively related to income and negatively related to hours worked, and that income is a measure of an entity's potential productivity. It was concluded that a linear income-tax schedule (with the possibility of imposing negative tax), is optimal. Such notions of optimality rely on a social welfare function (of individual utilities) and further assumptions on the shape of individual utility functions, such as concavity. While these assumptions are not necessarily unreasonable, the reliance of the arguments on utility theory and a particular social welfare function leaves any sceptic of this branch of utility theory unconvinced — first, of the justification of redistributive actions, and secondly, of the existence of some optimal amount of redistribution.

Representative agent models of increasing complexity have since been developed in order to interpret macroeconomic phenomena, make predictions and develop optimal tax structures suitable for application in reality. Examples of such models include [15, 24, 100, 153]. The focus of the literature on optimal redistribution therefore departed from its original pursuit of a theoretical justification of a basic premise (redistribution in general and progressive taxation in particular), to attempts at accurate modelling of complex socio-economic systems.

Stiglitz [138], among others, contributed significantly to interpreting these developments and was instrumental in directing attention to the importance of inequality by emphasizing its far reaching economic implications. Furman and Stiglitz [61] have stated that representative agent models can be misleading, and that neoclassical models which assume that distribution and efficiency are separable, are wrong in the context of the informational limitations of reality. Such objections are unsurprising. Given the complexity of the system under consideration, the extent of the simplifying assumptions required in order to build any model of explicit interaction (heterogeneous agents choosing between consumption and saving, for example), will invite critique, and be spurious. Furthermore, the utility framework on which most models of optimal redistribution relies, while perhaps theoretically acceptable, is notoriously difficult to capture empirically and apply in reality. Indeed, developing a practical theory for redistributive tax seems to be intractable.

This review suggests that the most fruitful theoretical investigations of redistribution have been aimed at theoretical justification, rather than practical estimates or predictions. Existing theoretical justifications of redistribution rely on assumptions related to utility and efficiency. This naturally leads to the question of whether there are simpler or other assumptions upon which

[1]An allocation of resources is called *Pareto efficient*, after the Italian engineer and economist Vilfredo Pareto, if no person can be made better off without rendering another worse off [28].

other justifications for redistributive actions can be constructed. In other words, can redistributive actions be justified in ways other than in terms of a specific version of utility theory and perhaps even over and above the idea that redistribution is required for social stability, as mentioned in §2.2?

2.5 Metrics of economic inequality

The predominant literature on the measurement of economic inequality distinguishes between the statistical problem of the measurement of the inequality of a given variable describing the distribution of some physical quantity, and the specific problem of measuring economic inequality [11, 46]. Economists are not concerned primarily with the distribution of wealth or income, but with the distribution of welfare (and the maximisation thereof) [46]. The use of any given measure of economic inequality implies certain assumptions about the function mapping the distribution of wealth to that of welfare, called the *social welfare function* [39]. The reason why certain implicit assumptions about a social welfare function are included in any measure of economic inequality is that different measures differently emphasize the extent of inequality at different points along the wealth spectrum.

Various requirements that should be satisfied by a good metric of economic inequality have been put forward. These differ in strength and some are even contradictory. When choosing a certain metric, its exact purpose should therefore be established carefully, and its characteristics noted explicitly.

Certain fundamental requirements are commonly accepted. The *anonymity principle* [64], for example, requires that the metric should depend on wealth only, and that permutations of entity labels should leave it invariant. The *population principle* furthermore states that the inequality metric should be independent of population size, in other words, all distributions obtained through the replication of a certain distribution should exhibit identical inequality [46]. According to the *principle of transfers*, any transfer of wealth d from an entity with wealth w_1 to another of wealth w_2, where $w_1 - d > w_2$, should furthermore result in a decrease in the inequality metric [46, 115]. Such a transfer is equivalent to a mean-preserving spread [11]. The *decomposability principle* finally requires that if two distributions W_1 and W_2 of the same size and mean, are merged with a third distribution W_3, then it should hold that $I(W_1, W_3) > I(W_2, W_3)$, if $I(W_1) > I(W_2)$, where I is the inequality measure.

Further requirements may be imposed, depending on the particular use of the metric. For example, conflicting viewpoints exist in respect of how an inequality metric should be influenced by proportional increases in wealth. During a proportional increase, the distributional shape remains unchanged, the absolute difference between entity wealth positions increases, and, if the social welfare function is increasing over wealth (as is typically assumed [22]), the total welfare obtained by the population is increased. These three observations may be used to qualify assertions that a good inequality metric should be uninfluenced, increasing, or decreasing, respectively, when subject to proportional increases in wealth.

These disputes may be resolved to some extent by distinguishing between increased equality and increased welfare, as is done in [39], but any single quantification of a distribution's inequality still contains within it an assumption of the weighting of inequality within different parts of the wealth spectrum. For example, the popular *Gini coefficient* weighs inequality in the tails of the distribution less heavily than in the middle [11], while *variance* weighs inequality farther from the mean more heavily. While it may be expected that mean deviation is an objective metric in the sense that it weighs no part of the wealth spectrum more heavily in the quantification of inequality, it remains uninfluenced by transfers between entities on the same side of the mean, and is therefore clearly a poor indication of the extent of inequality present in a given distribution [46].

The principle of *scale invariance* [39] states that if some distribution $W_1(y)$ is considered more equal than some other distribution $W_2(y)$, then for any constant k, $W_1(ky)$ should be more equal than $W_2(ky)$. The principle of *translation invariance* [38], however, states that if some distribution $W_1(y)$ is considered more equal than a distribution $W_2(y)$, then for any constant k, $W_1(y+k)$ should be more equal than $W_2(y+k)$. Except in trivial cases, these two principles cannot hold simultaneously.

Adopting the principle of scale invariance, together with the commonly accepted principles mentioned above, implies that a continuous inequality metric must be ordinally equivalent to

$$I_{GE}(y) = \frac{1}{s(s-1)}\left[\frac{1}{n}\sum_{i=1}^{n}\left[\frac{y_i}{\mu(y)}\right]^s - 1\right], \tag{2.1}$$

where μ is the mean wealth, distributed between n entities, and $s \in \mathbb{R}$ is a parameter which determines the metric's sensitivity to inequalities at different positions in the wealth distribution [38]. For positive or negative values of s, the metric is more sensitive to changes at the top or bottom of the wealth distribution, respectively. This class of measures is referred to as the class of *generalised entropy* measures [40], and is an important subclass of the class of *relative inequality measures* [22].

Choosing to adopt the principle of translation invariance instead of that of scale invariance gives rise to the class of continuous inequality indices

$$I_A(y) = \frac{1}{n}\sum_{i=1}^{n} e^{s[y_i-\mu(y)]} - 1, \tag{2.2}$$

when s, the sensitivity parameter, is nonzero, or

$$I_A(y) = \frac{1}{n}\sum_{i=1}^{n}[y_i - \mu(y)]^2 \tag{2.3}$$

when $s = 0$, which is simply the variance. These indices reside in the class of *absolute inequality measures* [21].

The most commonly used metric of economic inequality, the *Gini coefficient* [39], is defined by

$$I_{Gini}(y) = \frac{1}{2n^2\mu(y)}\sum_{i=1}^{n}\sum_{j=1}^{n}|y_i - y_j|. \tag{2.4}$$

This measure is similar to the relative measures, but requires that the decomposition principle be relaxed in the sense that the arbitrarily introduced distribution does not overlap with the initial distributions [38].

The most intuitive interpretation of the Gini coefficient is that it represents the area between a given distribution's Lorenz curve[2] and the line of total equality. The curve represents the proportion of total wealth belonging to a given bottom proportion of the population, as illustrated in Figure 2.1.

[2]The Lorenz curve is widely used in the calculation of inequality indices [64]. It is defined implicitly by

$$L(F) = \frac{1}{\mu}\int_0^{y_1} y f(y)\,\mathrm{d}y, \quad F = \int_0^{y_1} f(y)\,\mathrm{d}y, \tag{2.5}$$

where f is the density distribution of wealth.

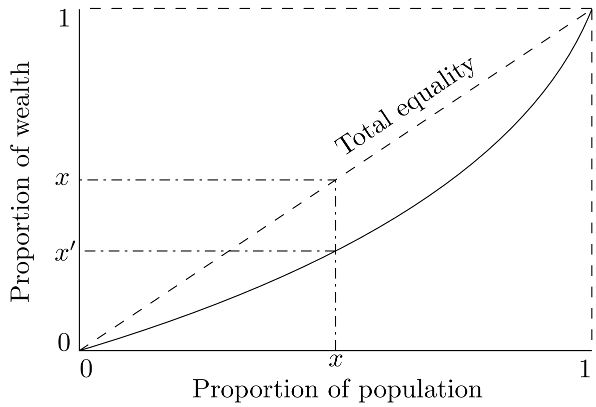


FIGURE 2.1: *The Lorenz curve depicts the proportion of total wealth x' owned by the bottom proportion x of the population which, in a completely equal population, is equal to x.*

The assumption that marginal economic welfare decreases as wealth position increases is commonly accepted, and leads to the conclusion that greater equality leads to greater total welfare [46]. Here, therefore, lies another motivation for seeking greater equality in wealth distribution, highlighting the importance of investigating various redistribution techniques. Finding a universally accepted social welfare function, however, proves difficult because of the diversity of human preference and opinion [130].

The focus in this thesis is on the effects that certain redistribution schemes have on a wealth distribution, and not primarily on the welfare resulting from the wealth distribution. The precise sensitivity weighting of a given inequality metric, which is related to the shape of the social welfare function, is therefore not of primary importance. The distinction between absolute and relative indices is, however, of fundamental importance, since the concern in this thesis is an analysis of changes in distributional shape, rather than absolute differences. Since the primary intention is to differentiate between upward or downward trends in the evolution of relative inequality, any metric of relative inequality will therefore suffice, since the exact choice only influences the rate of change, for finite sensitivity parameters.

2.6 Existing models of wealth distribution

If models describing the evolution of wealth distribution were arranged along a linear spectrum with models which mimic reality by modelling economic interactions explicitly (a so-called bottom-up approach) on the extreme left, and models which simply map distributional dynamics over time directly using functions which reproduce certain characteristics of observed distributional changes (a so-called top-down approach) on the extreme right, then the present investigation would fall towards the right-hand side of most existing models in the literature. The prevailing models of wealth distribution aim at explaining the driving forces behind the emergence of certain wealth distributions, subject to various growth and redistributive assumptions. The aim in this thesis, however, is the investigation of different redistributive effects upon a given wealth distribution, given that certain inherent characteristics of wealth-distribution dynamics, and specifically increases in inequality, are present.

This distinction is the reason for the adoption of a greater degree of abstraction, or expressions farther removed from reality, in the present work. Another implication is that a review of existing models will contribute little to the argument of this thesis. A brief review of the existing models is nevertheless included to establish context.

Models of wealth distribution may be categorised broadly into two categories. The aim in the first category is to find functional descriptions of the shape of empirically observed wealth distributions, so that, once a certain parameter set has been estimated, reasonable approximations can be obtained for the distribution shape and the level of inequality. Examples of such models include the Pareto models [109, 110, 111], the lognormal model [43], the two-parameter logistic model [33, 58], the Pearson Type V model [12, 147] and the Dagum models [44, 45].

The aim in the second category is to reproduce certain distribution shapes (characterised by the distribution functions of the first category) in the steady state solutions of the models by modelling economic interactions over time, either in continuous or discretised spaces.

In the model of Hugget [75], Aiyagari [6] and Bewley [19], the wealth distribution of a continuum of infinite households is, for instance, modelled over discrete time. The evolution of household income is exogenously described, and households choose between consumption and saving by maximising a given utility function. In the treatment by Hugget [75], each agent receives an endowment of one perishable consumption good during each of an infinite number of successive time periods, which is either high or low, denoted by e_h or e_ℓ, respectively. An individual agent's endowment follows a Markov process with stationary transition probabilities $\pi(e'|e)$ which are independent of all other agents' current or historical endowments. Each agent's preferences for consumption is defined by the utility function

$$E\left[\sum_{t=0}^{\infty}\beta^t u(c_t)\right],\ \beta\in(0,1),$$

where E denotes expectation, $u(c_t)$ is the utility derived from consumption c_t during perdiod t, β is the utility discount factor and

$$u(c)=\frac{c^{1-\sigma}}{1-\sigma},\ \sigma>1.$$

Here the parameter σ determines the rate of diminishing marginal utility.

An entity has access to a credit balance a which entitles the agent to a goods during the current period. By paying $a'q$ during the current period, an agent gains a' units for the following period, where q is the price of next-period credit balances. A limit is imposed on credit balances, denoted by $a > \underline{\mathrm{a}}$ where $\underline{\mathrm{a}} < 0$.

Each agent's decision problem is then

$$v(x;q)=\max_{(c,a')\in\Gamma(x;q)} u(c)+\beta\sum_{e'} v(a',e';q)\pi(e'|e),$$

where the period budget constraint is given by

$$\Gamma(x;q)=\{(c,a'):c+a'q\le a+e,\ c\ge 0,\ a'\ge \underline{\mathrm{a}}\}.$$

This model has also been translated to the continuous time domain for greater ease of analysis [2].

The great mathematician Benoit Mandelbrot, drew the attention of economists to the principles of statistical mechanics (such as the mathematical description of the aggregate movement of colliding gas particles), by highlighting similarities between the fields of statistical physics and macro economics [97]. A vast body of literature has since emerged, building on such similarities, within a field now referred to as *econophysics* [52]. In models within this realm, economic agents are often likened to individual gas particles. Collisions between agents constitute possible transactions between them which follow certain trade rules. These assumptions give rise to a spatially homogeneous Boltzman equation[3] [52] which models economic characteristics of a continuum of agents. The effects of different wealth conservation assumptions, based on the types of market risk included, or of different saving propensities between agents, have, for example, been investigated [35].

Models originating from biological mathematics, and specifically the well-known Lotka-Volterra models[4] of population growth and interspecies interaction, have also inspired different models of wealth distribution within the field of economics [20, 96, 128, 135]. Models aimed specifically at reproducing certain distribution shapes, and modelling their evolution over time, typically have a finite number of agents whose individual wealth levels evolve randomly (according to some assumed distribution), but with each receiving a minimum wealth dependent on the average population wealth, and with competition limiting the growth of the wealthiest individuals, based on their individual wealth levels and the population average wealth level [96]. Such models approach generalised Lotka-Volterra models for large numbers of agents, and have been shown to produce Pareto distributions [134].

Inspiration has therefore been drawn from utility theory, physics and biological mathematics, in the pursuit of describing the distribution of wealth in an economy. Certain reservations pertaining to these models, and econometric models in general, have also been voiced [48, 56, 87, 102]. For example, the belief that empirical regularities should exist in several areas of economic activity, and that certain shortcomings are inherent to modelling transactions, leading to the confusion of basic concepts, have been criticised [63].

Many of these criticisms stem from the apparent inability of proposed models to produce accurate predictions [102]. The purpose of modelling need not necessarily be prediction, however — it may be the illumination of core dynamics, a search for new questions and/or the promotion of a scientific habit of mind [53].

The aim of wealth distribution models encountered in the literature is therefore to describe and reproduce empirically observed distribution shapes. The models pursued in this thesis, on the other hand, are related to economic reality only in the sense of allowing the increase of relative inequality over time — no other direct relation to existing or historic wealth distributions is sought. The aim is simply to derive high-level conceptual insights, thus avoiding direct relations between empirical controversies and refraining from attempts to model interactions explicitly.

2.7 Chapter summary

The discourse on economic inequality was reviewed very briefly in this chapter from the point of view of social justice, and the recurring theme of limited egalitarianism in terms of resource ownership was identified. Empirically observed historic trends in economic inequality, and certain resulting postulates, were then described. The notion of redistribution of wealth was briefly reviewed by discussing various means of redistribution, the desirability thereof, and its relationship to economic inequality. The notion of optimal redistribution and the theoretical justification of redistributive acts were then considered. The measurement of economic inequality was discussed

[3] The spatially homogeneous Boltzman equation [104] is defined by $\partial w/\partial t = Q(w)$ where w is a time-dependent probability density for the velocities of particles of a dilute gas, and Q is a collision operator which captures the dynamics of colliding particle interactions according to certain collision assumptions.

[4] Lotka-Volterra models are systems of coupled ordinary differential equations which describe the evolution of interacting species' population sizes over time, such as the predator-prey model [42] $\mathrm{d}w_1/\mathrm{d}t = aw_1 - bw_2$, $\mathrm{d}w_2/\mathrm{d}t = cw_1w_2 - dw_2$, where w_1 and w_2 denote the prey and predator population densities at time t, respectively, and a, b, c and d denote constants of proportionality.

next and two important classes of indices for measuring inequality were reviewed. Existing models of wealth distribution were finally reviewed briefly and the aim of this thesis was elucidated within the context of the literature review presented in this chapter.

CHAPTER 3

Mathematical prerequisites

Contents

Certain mathematical definitions, theorems and techniques required to follow the arguments presented later in this thesis are reviewed in this chapter. First, basic definitions and theorems often used in the calculus of functions of real variables and in the proofs presented later in this thesis, are reviewed. A brief background is then provided on differential equations and their solution properties, and this is followed by a description of a well-known classification of second-order linear PDEs. Examples of equations from a specific subclass of these PDEs, namely the class of reaction-diffusion equations, are then provided, and this is followed by a review of an existence and uniqueness result for initial-boundary value problems involving reaction-diffusion equations. A method for determining the stability properties of equilibrium solutions to certain differential equations is reviewed next. Finally, numerical techniques that are utilised to approximate solutions to differential equations are noted and the class of finite difference methods, in particular, is described, as these methods are applied extensively later in this thesis.

3.1 Mathematical definitions

In this section, certain prerequisite mathematical definitions are provided which will prove useful later in the exposition of this thesis. The definition of a vector space is the chosen point of departure. First, the notions of *addition*, *scalar multiplication*, and *equality* are, however, considered in the context of vectors. Recall that a *vector of dimension* n is an ordered set of n numbers.

Definition 3.1 *(Addition, scalar multiplication and equality [152, p. 298])*
Let $\boldsymbol{a} = [a_1, a_2, \ldots, a_n]$ *and* $\boldsymbol{b} = [b_1, b_2, \ldots, b_n]$ *be two vectors of length* $n \in \mathbb{N}$*. Then*

(i) $\boldsymbol{a} + \boldsymbol{b} = [a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n]$ *(addition),*

(ii) $k\boldsymbol{a} = [ka_1, ka_2, \ldots, ka_n]$, *for each* $k \in \mathbb{R}$ *(scalar multiplication), and*

(iii) $\boldsymbol{a} = \boldsymbol{b}$ *if and only if* $a_1 = b_1, a_2, = b_2, \ldots, a_n = b_n$ *(equality).*

The notion of a vector space may be defined in terms of the aforementioned operations.

Definition 3.2 *(Vector space [152, p. 327])*
Let $\mathcal{V}$ *be a set of elements on which the operations of vector addition and scalar multiplication are defined, and let* $\boldsymbol{a}, \boldsymbol{b}$ *and* $\boldsymbol{c}$ *be any elements of* $\mathcal{V}$. *Then* $\mathcal{V}$ *is a vector space if, for vector addition, it holds that*

(i) $\boldsymbol{a} + \boldsymbol{b} \in \mathcal{V}$,

(ii) $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$,

(iii) $(\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c})$,

(iv) there is a unique vector $\boldsymbol{0} \in \mathcal{V}$ *such that* $\boldsymbol{0} + \boldsymbol{a} = \boldsymbol{a} + \boldsymbol{0} = \boldsymbol{a}$,

(v) there exists a vector $-\boldsymbol{a} \in \mathcal{V}$ *such that* $\boldsymbol{a} + (-\boldsymbol{a}) = (-\boldsymbol{a}) + \boldsymbol{a} = \boldsymbol{0}$,

and, for scalar multiplication, it holds that

(vi) $k_1\boldsymbol{a} \in \mathcal{V}$ *for each* $k_1 \in \mathbb{R}$,

(vii) $k_1(\boldsymbol{a} + \boldsymbol{b}) = k_1\boldsymbol{a} + k_1\boldsymbol{b}$ *for each* $k_1 \in \mathbb{R}$,

(viii) $(k_1 + k_2)\boldsymbol{a} = k_1\boldsymbol{a} + k_2\boldsymbol{a}$ *for each pair* $k_1, k_2 \in \mathbb{R}$,

(ix) $k_1(k_2\boldsymbol{a}) = (k_1k_2)\boldsymbol{a}$ *for each pair* $k_1, k_2 \in \mathbb{R}$,

(x) $1\boldsymbol{a} = \boldsymbol{a}$.

It is necessary to be able to express the relative size of a vector, or more importantly, the size of the difference between vectors, in order to define notions such as continuity and convergence. Functions called *norms* are employed for this purpose.

Definition 3.3 *(A norm on a vector space [92, p. 121])*
A real-valued function p *defined on a vector space* $\mathcal{V}$ *is called a norm if*

(i) $p(\boldsymbol{a}) > 0$ *whenever* $\boldsymbol{a} \neq \boldsymbol{0}$,

(ii) $p(k_1\boldsymbol{a}) = |k_1|p(\boldsymbol{a})$ *for any* $k_1 \in \mathbb{R}$, *and*

(iii) $p(\boldsymbol{a} + \boldsymbol{b}) \leq p(\boldsymbol{a}) + p(\boldsymbol{b})$,

for all $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{V}$.

If p is a norm defined on a vector space $\mathcal{V}$ and $\boldsymbol{a} \in \mathcal{V}$, then it is common to denote the value of $p(\boldsymbol{a})$ by $\|\boldsymbol{a}\|$. The most commonly used norms in $\mathbb{R}^n$ are $\|\boldsymbol{a}\|_1 = \sum_{i=1}^{n} |a_i|$, the *Euclidean norm*

$$\|\boldsymbol{a}\|_2 = \left(\sum_{i=1}^{n} a_i^2 \right)^{1/2},$$

and $\|\boldsymbol{a}\|_\infty = \max_i\{|a_i|\}$ [92]. The vector spaces employed in this thesis include the infinite dimensional vector space of continuous real-valued functions defined on a real interval, say $[\lambda_a, \lambda_b]$. The most commonly used norms (according to [92]) on this particular vector space are the 1-*norm*

$$\|h\|_1 = \int_{\lambda_a}^{\lambda_b} |h(\lambda)| \, \mathrm{d}\lambda, \tag{3.1}$$

the 2-*norm*

$$\|h\|_2 = \left(\int_{\lambda_a}^{\lambda_b} |h(\lambda)|^2 \, \mathrm{d}\lambda \right)^{\frac{1}{2}}, \tag{3.2}$$

and the *uniform norm*

$$\|h\|_\infty = \max\{|h(\lambda)| : \lambda_a \leq \lambda \leq \lambda_b\}, \tag{3.3}$$

where h is a continuous real-valued function defined on the interval $[\lambda_a, \lambda_b]$. These are special cases of the p-norm

$$\|h\|_p = \left(\int_{\mathcal{V}} |h|^p \, \mathrm{d}u \right)^{\frac{1}{p}},$$

where h is a real-valued function defined on a vector space $\mathcal{V}$ and u is a measure of $\mathcal{V}$. The set of all real-valued functions with finite p-norms on a given region is denoted by the symbol L^p [143]. More specifically, the set $L^p(\mathcal{V}, \mathcal{D})$ denotes all functions defined on $\mathcal{V}$ mapping to $\mathcal{D}$ with finite p-norms (the limit as $p \to \infty$ is also allowed). The space L^∞ on a region of interest therefore includes all real-valued functions for which the integrals of the functions over the region are defined and the functions are bounded on the region.

The prerequisites for defining the notion of a limit in the context of multivariate functions have now been reviewed.

Definition 3.4 *(Limits for functions of multiple variables [92, p. 126])*
If a real-valued function w is defined on a subset $\mathcal{V}$ of $\mathbb{R}^n$, then the limit $\lim_{\boldsymbol{a}\to\boldsymbol{b}} w(\boldsymbol{a}) = L$ means that there exists, for any $\epsilon > 0$, a corresponding $\delta > 0$ such that $|w(\boldsymbol{a}) - L| < \epsilon$ if $\boldsymbol{a} \in \mathcal{V}$ and $0 < \|\boldsymbol{b} - \boldsymbol{a}\| < \delta$.

A function is continuous at a point if its limit exists there, the function itself is defined at the point, and the function and its limit at the point are equal. This notion is made more precise in the following definition.

Definition 3.5 *(Continuity of a real-valued function on a vector space [92, p. 126])*
Let h be a real-valued function defined on a vector space $\mathcal{V}$ on which a norm $\|\cdot\|$ is defined. The function h is called continuous at a point $\boldsymbol{a} \in \mathcal{V}$ if there exists a $\delta > 0$ for every $\epsilon > 0$ such that $\|\boldsymbol{b} - \boldsymbol{a}\| < \delta$ implies that $|h(\boldsymbol{b}) - h(\boldsymbol{a})| < \epsilon$.

If a function is continuous at all points in some set, then it is called continuous on that set [127]. A strong type of continuity, often used in the proofs of existence and uniqueness results for solutions to *ordinary differential equations* (ODEs) and PDEs, known as Lipschitz continuity, is defined as follows.

Definition 3.6 *(Lipschitz continuity [54, p. 205])*
A real-valued function h is Lipschitz continuous on a vector space $\mathcal{V}$ on which a norm $\|\cdot\|$ is defined if there exists a positive constant L_h so that

$$|h(\boldsymbol{b}) - h(\boldsymbol{a})| \leq L_h \|\boldsymbol{b} - \boldsymbol{a}\|, \tag{3.4}$$

for all $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{V}$.

It is important to note that a function may be locally Lipschitz continuous, yet not globally Lipschitz continuous. For example, the function $h(\lambda) = \lambda^2$ is not Lipschitz continuous on $\mathbb{R}$, but it is Lipschitz continuous on any bounded interval $[\lambda_a, \lambda_b] \in \mathbb{R}$.

The notion of a *derivative* plays an important role in the models put forward later in this thesis. Before the concept of a derivative can be made precise, however, the notion of a linear function is required.

Definition 3.7 *(Linear function [127, pp. 296–297])*
A real-valued function ℓ defined on a vector space $\mathcal{V}$ is linear if

$$\ell(k_1\boldsymbol{a} + k_2\boldsymbol{b}) = k_1\ell(\boldsymbol{a}) + k_2\ell(\boldsymbol{b}),$$

for all $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{V}$ and $k_1, k_2 \in \mathbb{R}$.

Definition 3.8 *(The derivative of a real-valued function [127, p. 312])*
If h is a real-valued function on a vector space $\mathcal{V}$ and $\boldsymbol{a}$ is an interior point of $\mathcal{V}$, then h is differentiable if there exists a linear functional ℓ on $\mathcal{V}$ with the property that there exists a $\delta > 0$ for every $\epsilon > 0$ where $\|\boldsymbol{a} - \boldsymbol{c}\| < \delta$ implies $\boldsymbol{c} \subset \mathcal{V}$ such that

$$|h(\boldsymbol{b}) - h(\boldsymbol{a}) - \ell(\boldsymbol{b} - \boldsymbol{a})| \leq \epsilon\|\boldsymbol{b} - \boldsymbol{a}\|.$$

The functional ℓ is called the derivative of h at $\boldsymbol{a}$ and is denoted by $h'(\boldsymbol{a})$. If the derivative of h exists at $\boldsymbol{a}$, the function h is said to be differentiable at $\boldsymbol{a}$.

The mean value theorem for a real-valued function will be required later in the derivation of numerical approximation schemes for differential equations.

Theorem 3.1 *(Mean value theorem [127, p. 344])*
Let h be a real-valued function on a vector space $\mathcal{V}$ and let $\boldsymbol{a}, \boldsymbol{b}$ be interior points in $\mathcal{V}$. If $(1-\alpha)\boldsymbol{a} + \alpha\boldsymbol{b}$ are interior points of $\mathcal{V}$ for all $\alpha \in (0, 1)$ and if h is differentiable at all these points, then there exists a point $\boldsymbol{c} = (1-\alpha_0)\boldsymbol{a} + \alpha_0\boldsymbol{b}$ for some $\alpha_0 \in (0, 1)$ such that

$$h(\boldsymbol{b}) - h(\boldsymbol{a}) = h'(\boldsymbol{c})(\boldsymbol{b} - \boldsymbol{a}).$$

If a derivative of a real-valued function of more than one variable is taken with respect to one of these variables, a *partial derivative* is obtained.

Definition 3.9 *(Partial derivative [78, pp. 395–396])*
If h is a real-valued function defined on a vector space $\mathcal{V}$ then for $\boldsymbol{a} \in \mathcal{V}$ the partial derivative with respect to the m-th entry of $\boldsymbol{a}$ is defined as

$$\frac{\partial h}{\partial a_m} = \lim_{\delta \to 0} \frac{h(a_1, \ldots, a_m + \delta, \ldots, a_n) - h(a_1, \ldots, a_m, \ldots, a_n)}{\delta}.$$

A function is called *differentiable* on a set if its derivative exists everywhere in the set. If the derivative of a function is continuous, the function is called *continuously differentiable*. The class of all continuous functions on a given region of interest is denoted by $\mathcal{C}^0$, with $\mathcal{C}^m$ representing the set of all m-times continuously differentiable functions on the region.

The functions to be considered in this thesis are real-valued functions of two variables. The relationship between the integral and the partial derivative of a continuous function of two variables is given by the following theorem.

Theorem 3.2 *(The fundamental theorem of the calculus [136, pp. 312–315])*
If $w(\lambda, t)$ is a continuous real-valued function on $[\lambda_a, \lambda_b] \times [t_a, t_b]$, then the function h defined by

$$h(\lambda, t) = \int_{\lambda_a}^{\lambda} w(v, \lambda)\,\mathrm{d}v, \;\; \lambda_a \leq \lambda \leq \lambda_b, \;\; t_a \leq t \leq t_b$$

is continuous on $[\lambda_a, \lambda_b] \times [t_a, t_b]$, the partial derivative of h with respect to λ exists on $(\lambda_a, \lambda_b) \times [t_a, t_b]$ and is given by $\partial h/\partial\lambda = w(\lambda, t)$.

Furthermore,

$$\int_{\lambda_a}^{\lambda_b} w(\lambda,t)\,\mathrm{d}\lambda = h(\lambda_b,t) - h(\lambda_a,t).$$

When one variable in a real-valued function of two variables is fixed, a real-valued function of a single real variable is obtained. Points where the first derivative of such a function evaluate to zero, or does not exist, are called *critical points* [136].

Theorem 3.3 *(Second derivatives at critical points [136, pp. 198–202])*
Let $h(\lambda)$ be a continuously differentiable real-valued function on a real interval $\mathcal{D}$. Then the second-derivative of h is negative (positive) at critical points with zero first derivatives which are local maxima (minima).

A theorem commonly used in the computational approximation of function values, known as Taylor's Theorem [30], is considered next. This theorem is presented here specifically in the form required to approximate a real-valued function of two variables in the direction of one of the variables.

Theorem 3.4 *(Taylor's theorem [30, pp. 10–11])*
For any continuous, $n+1$ times differentiable real-valued function w on $[\lambda_a,\lambda_b]\times[t_a,t_b]$, and $\lambda, \lambda+\delta \in [\lambda_a,\lambda_b]$, $t\in[t_a,t_b]$,

$$w(\lambda,t+\delta) = w(\lambda,t)+\delta\frac{\partial w(\lambda,t)}{\partial t}+\frac{\delta^2}{2!}\frac{\partial^2 w(\lambda,t)}{\partial t^2}+\cdots+\frac{\delta^{(n)}}{(n)!}\frac{\partial^n w(\lambda,t)}{\partial t^n}+\frac{\delta^{(n+1)}}{(n+1)!}\frac{\partial w^{(n+1)}(\lambda,t)}{\partial t^{n+1}}(t+\alpha\delta)$$

for some $\alpha\in[0,1]$.

When function approximations are made, the size of the error is often denoted using so-called *Big O notation*. This notation is defined as follows.

Definition 3.10 *(Big O notation [13])*
Let h and w be real-valued functions defined on an interval of real numbers $[\lambda_a,\lambda_b]$. Then the notation $h(\lambda)=\mathcal{O}(w(\lambda))$ means that for all $\lambda\in[\lambda_a,\lambda_b]$ there exists a positive constant k_1 such that $|h(\lambda)|\le k_1|w(\lambda)|$.

An important relationship between the integral of a function squared and the square of the integral of the function will be of interest later in this thesis. With a view to establishing such a relationship, a theorem on inner products within real vector spaces is quoted. First, however, the notion of the *inner product* of two vectors is recounted.

Definition 3.11 *(The inner product of two vectors [136, p. 824])*
Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two vectors of length n, then the inner product of $\boldsymbol{a}$ and $\boldsymbol{b}$ is

$$\boldsymbol{a}\cdot\boldsymbol{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

Theorem 3.5 *(Cauchy-Buniakovskii-Schwarz inequality [127, p. 202])*
If $\mathcal{V}$ is a real vector space and $\cdot$ is an inner product defined on $\mathcal{V}$, then

$$|\boldsymbol{h}\cdot\boldsymbol{g}| \le \sqrt{(\boldsymbol{h}\cdot\boldsymbol{h})(\boldsymbol{g}\cdot\boldsymbol{g})},$$

for any $\boldsymbol{h}$, $\boldsymbol{g}\in\mathcal{V}$.

An inner product may be defined on the space of continuous real-valued functions by

$$h\cdot g = \int_{\lambda_a}^{\lambda_b} hg.$$

It therefore follows that the integral of a real-valued function squared is at least as large as the square of the integral of the function. This inequality is stated mathematically as follows.

Corollary 3.1 *If h is a real-valued function of a real variable on the interval $[\lambda_a, \lambda_b]$, then*

$$\left[\int_{\lambda_a}^{\lambda_b} h(\lambda)\,\mathrm{d}\lambda\right]^2 \leq \int_{\lambda_a}^{\lambda_b} h^2(\lambda)\,\mathrm{d}\lambda.$$

3.2 Differential equations

A very brief overview of the extensive field of differential equations is provided in this section. The primary focus is placed on concepts and theorems that will be of use at a later stage in this thesis. This particular need dictates the composition of this section.

A mathematical equation defines a relation between a given set of variables. A differential equation is an equation which contains the derivatives of at least one function, often called the dependent variable, with respect to at least one independent variable [152].

If the derivatives in such an equation are all ordinary derivatives with respect to the same independent variable, the equation is called an ODE. When partial derivatives with respect to at least two independent variables are, however, present, the equation is called a PDE. A differential equation is considered linear if it contains only linear combinations of the dependent variable and its derivatives, with the coefficients being either constants or functions of the independent variables only [23]. The order of a differential equation, ordinary or partial, is the highest order derivative present in the equation.

An n-th order linear ODE is therefore an equation of the form

$$\xi_n(t)\frac{\mathrm{d}^n w}{\mathrm{d}t^n} + \cdots + \xi_2(t)\frac{\mathrm{d}^2 w}{\mathrm{d}t^2} + \xi_1(t)\frac{\mathrm{d}w}{\mathrm{d}t} + \xi_0(t)w = h(t), \tag{3.5}$$

where ξ_n is not identically zero, and $\xi_i = \xi_i(t)$ is a (possibly constant) function of the independent variable t for all $i \in \{0, 1, \ldots, n\}$. A second order linear PDE involving two independent variables is an equation of the form

$$\xi_1(\lambda,t)\frac{\partial^2 w}{\partial\lambda^2} + \xi_2(\lambda,t)\frac{\partial^2 w}{\partial\lambda\partial t} + \xi_3(\lambda,t)\frac{\partial^2 w}{\partial t^2} + \xi_4(\lambda,t)\frac{\partial w}{\partial\lambda} + \xi_5(\lambda,t)\frac{\partial w}{\partial t} + \xi_6(\lambda,t)w = h(\lambda,t). \tag{3.6}$$

If h is identically zero in (3.5) or (3.6), these equations are said to be *homogeneous* [152].

A solution to a differential equation on a given region is any function which, if substituted into the equation, reduces it to an identity on that region. If a solution to an n-th order differential equation is n times continuously differentiable, the solution is called a *classical solution*, as opposed to a *weak solution* for which the derivatives of order n or smaller may not all exist, but which nonetheless satisfies the differential equation in some predefined sense [55]. For example, in applications, PDEs are sometimes solved on irregularly shaped domains (consider the internal strain in an irregularly shaped mechanical component subject to various forces), which do not necessarily have smooth boundaries, and so function derivatives may therefore not be defined everywhere, yet the quantity modelled by the dependent variable necessarily has a solution on the domain.

If a solution is defined on a real interval $[t_0, t_a)$ and a solution exists on $[t_0, t_a+\delta)$, for some $\delta > 0$, which coincides with the solution on $[t_0, t_a)$, then the solution on $[t_0, t_a)$ is called *continuable* [69]. A solution which is not continuable is called *noncontinuable*.

Another way of defining the solution to a differential equation is as the primitive which gives rise to the differential equation. A *primitive* is an equation, involving n essentially arbitrary constants, which establishes a relation between variables [14]. Using differentiation and back-substitution, a primitive may be transformed into a differential equation. Substituting a primitive into its corresponding differential equation therefore gives rise to an identity, as required by the first definition.

Not all differential equations can be solved analytically [133]. This may not only be because ingenuity fails, but may also be because the solutions to some differential equations cannot be expressed in terms of standard functions. Analytic methods exist for obtaining solutions to certain classes of differential equations. For example, linear ODEs of the form (3.5) with constant coefficients $\xi_0, \ldots, \xi_n$ may be solved by means of the substitution $w = e^{mt}$ (where t is the independent variable) so as to obtain a polynomial equation in m. The roots of the polynomial equation may then be used to construct a general solution to the ODE (see, for example, [152]). If it is expected that a solution to a PDE in two independent variables is a product of two functions of single variables, the solution may be obtained by separating these functions to different sides of the equation, thus obtaining two ODEs which may be solved separately to find a general solution to the PDE [23].

For a solution to a differential equation to be unique, however, the values of all the essentially arbitrary constants in the primitive need to be fixed. This may be achieved by specifying conditions, such as a set of initial values, that the solution to an equation of the form (3.5) must satisfy at some value $t = t_a$. The equation (3.5) together with these *initial conditions* is collectively referred to as an *initial value problem*. An initial-value problem is considered *well posed* if a unique solution to the problem exists and this solution depends continuously on the initial data [31]. For example, a first order ODE of the form

$$\frac{\mathrm{d}w}{\mathrm{d}t} = h(w, t) \tag{3.7}$$

requires a single initial condition, according to the following theorem, in order to admit a unique solution.

Theorem 3.6 *(Uniqueness of solutions to first order ODE initial-value problems [152])*
Let $\mathcal{R}$ be a rectangular region in t and w that contains a point (t_0, w_0). If h and $\partial h/\partial w$ are continuous in $\mathcal{R}$, then there exits an interval $\mathcal{I}$ centred on t_0, in which a unique function $w(t)$ satisfies (3.7).

It can be shown that a solution to the initial-value problem

$$\frac{\mathrm{d}w}{\mathrm{d}t} = h(w, t) \tag{3.8}$$
$$w(t_0) = w_0 \tag{3.9}$$

is bounded on a small interval containing t_0 if the function h is continuous.

Theorem 3.7 *(Boundedness of the solution to* (3.8)–(3.9) *[27])*
Any solution to the initial-value problem (3.8)–(3.9) *is bounded on* $[t_0, t_0 + T_{\max}]$*, for some* $T_{\max} > 0$*, if h is continuous on* $[t_0, t_0 + T_{\max}]$.

Similar results pertaining to the existence, uniqueness and boundedness of solutions to initial-boundary value problems involving PDEs are considerably more complex than in the case of ODEs. The solution techniques and theorems on properties of solutions applicable to PDEs also differ depending on the specific class of PDEs under consideration. Second-order linear PDEs may be classified into four distinct classes [23]. Consider the second-order linear PDE in two dependent variables,

$$\xi_1 \frac{\partial^2 w}{\partial \lambda^2} + \xi_2 \frac{\partial^2 w}{\partial \lambda \partial t} + \xi_3 \frac{\partial^2 w}{\partial t^2} + \xi_4 \frac{\partial w}{\partial \lambda} + \xi_5 \frac{\partial w}{\partial t} + \xi_6 w = h(\lambda, t), \tag{3.10}$$

where w and h are continuous functions of two real variables λ and t, ξ_j is a constant for all $j \in \{1, 2, 3, 4, 5, 6\}$ and $\xi_1^2 + \xi_2^2 + \xi_3^2 > 0$. There exists a transformation of variables

$$\begin{aligned} x &= k_1\lambda + k_2 t, \\ y &= -k_2\lambda + k_1 t, \text{ and} \\ W(x, y) &= p^{-1} e^{k_3\lambda + k_4 t} w(\lambda, t), \end{aligned}$$

where k_1, k_2, k_3, k_4 and $p \neq 0$ are constants with $k_1^2 + k_2^2 = 1$, so that (3.10) takes the form of one of the following cases [23]:

(a) if $\xi_2^2 - 4\xi_1\xi_3 > 0$, the equation is called *hyperbolic*,

(b) if $\xi_2^2 - 4\xi_1\xi_3 < 0$, the equation is called *elliptic*,

(c) if $\xi_2^2 - 4\xi_1\xi_3 = 0$ and $2\xi_3\xi_4 \neq \xi_2\xi_5$ or $2\xi_1\xi_5 \neq \xi_2\xi_4$, the equation is called *parabolic*,

(d) if $\xi_2^2 - 4\xi_1\xi_3 = 0$, $2\xi_3\xi_4 = \xi_2\xi_5$ and $2\xi_1\xi_5 = \xi_2\xi_4$, the equation is called *degenerate*.

The PDEs to be considered in this thesis are of the parabolic type. Parabolic PDEs typically involve a first order derivative with respect to a single real variable representing time (the independent variable), and second order derivatives with respect to a set of real variables modelling space [133]. The simplest example from this class is the so-called *heat equation*

$$\frac{\partial w}{\partial t} = k_d \frac{\partial^2 w}{\partial \lambda^2}, \tag{3.11}$$

where k_d is a constant denoting the diffusion rate. The heat equation models the spatial spreading of heat in a homogeneous material in one spatial dimension $\lambda \in [\lambda_a, \lambda_b]$ over time t.

3.3 Reaction-diffusion systems

A class parabolic PDEs, known as reaction-diffusion equations, contains equations of the form [84]

$$\frac{\partial \boldsymbol{w}}{\partial t} = \boldsymbol{g}(\boldsymbol{w}) + \boldsymbol{D}\nabla\boldsymbol{w}, \tag{3.12}$$

where ∇ denotes the Laplacian operator[1] and where $\boldsymbol{w} = \boldsymbol{w}(\boldsymbol{\lambda}, t)$ is an m-column vector of real-valued functions of time $t \in [0, \infty)$ on some spatial domain $\boldsymbol{\lambda} \in \mathcal{D}$. The second term on the right-hand side of (3.12), called the *diffusion term*, models the movement of components of $\boldsymbol{w}$ from areas of high density to areas of lower density. The matrix $\boldsymbol{D}$ is an $m \times m$ diagonal matrix containing a diffusion rate d_i as its i-th entry, for all $i \in \{1, 2, \ldots, m\}$. The first term on the right-hand side of (3.12) is called the *reaction term*, and it models the growth or decline of the components of $\boldsymbol{w}$. In general, this growth or decline may depend on the different component densities, the position in space, and time, such that $g = g(\boldsymbol{w}, t, \boldsymbol{\lambda})$. In all the cases considered in this thesis, however, the equation form (3.12) is sufficient.

[1]The Laplacian operator is a differential operator defined as $\nabla \equiv \sum_{i=1}^{n} \frac{\partial^2}{\partial \lambda_i^2}$ on $\mathbb{R}^n$ [23, p. 341].

Reaction-diffusion equations arise in a variety of applications. The spatio-temporal evolution of chemical reactions [146], or of biological species population densities [32], are two notable examples. The simplest instance of such an equation is the single component case, known as the *Kolmogorov-Petrovskii-Piskounov equation* [93]. *Fisher's equation* [57],

$$\frac{\partial w}{\partial t} = w(1-w) + d\frac{\partial^2 w}{\partial \lambda^2}, \tag{3.13}$$

proposed as a model for the advance of an advantageous gene in a population of density $w = w(\lambda, t)$ at time t and at position λ within a one-dimensional spatial domain, is an example from this class.

An excellent illustration of how equations of mathematically similar form may arise in unrelated fields is found here. If the reaction term in (3.13) is modified to $w(1-w^2)$, the so-called *Newell-Whitehead-Segel equation* [105] is obtained, which models convection in a plane of liquid which is heated from below.

As mentioned in §3.2, certain conditions are to be specified in order to guarantee that a solution to a differential equation is unique, if it exists. Reaction-diffusion equations model the spatio-temporal evolution of certain component densities, and for a unique solution to exist, the characteristics of the solution on the spatial domain's boundaries must be specified, as well as the initial density distributions. These conditions are referred to as the boundary- and initial-conditions, respectively. A parabolic PDE, together with initial and boundary conditions, is referred to as an *initial-boundary value problem.*

The initial condition is simply a function or set of functions defined on the spatial domain which maps each position in space to $\mathbb{R}^n$, corresponding to the initial density of each of n components. For some domain $\mathcal{D}$ with boundary $\partial\mathcal{D}$, linear boundary conditions specify a linear relationship between the dependent variable w and its partial derivatives of lesser order than that of the differential equation considered [76]. For a second order PDE, several types of boundary conditions may therefore be imposed. If the dependent variable value w at the boundary is specified, the boundary conditions are called *Dirichlet conditions.* If the derivative of w with respect to the normal direction on the boundary is specified, the boundary conditions are referred to as *Neuman conditions.* When the value of the derivative of w with respect to the unit normal on the boundary, as well as the function value w, is specified on the boundary, the boundary conditions are called *Cauchy boundary conditions.* A weighted combination of Neuman and Dirichlet conditions is called *Robin boundary conditions.* It is also possible to specify different boundary conditions on different subsets of the domain boundary, and conditions of this is type are referred to as *mixed boundary conditions.*

Existence and uniqueness of solutions to certain classes of initial-boundary value problems involving reaction-diffusion systems have been established in the literature. Morgan [103], for example, considered m-component systems of the form

$$\text{D.E.} \quad \frac{\partial \boldsymbol{w}}{\partial t} = \boldsymbol{g}(\boldsymbol{w}(\boldsymbol{\lambda}, t)) + \boldsymbol{D}\nabla \boldsymbol{w}(\boldsymbol{\lambda}, t),\ \boldsymbol{\lambda} \in \mathcal{D},\ t > 0, \tag{3.14}$$

$$\text{B.C.} \quad \boldsymbol{B}\boldsymbol{w}(\boldsymbol{\lambda}, t) = 0,\ \boldsymbol{\lambda} \in \partial\mathcal{D},\ t > 0, \tag{3.15}$$

$$\text{I.C.} \quad \boldsymbol{w}(\boldsymbol{\lambda}, 0) = \boldsymbol{w}_0(\boldsymbol{\lambda}),\ \boldsymbol{\lambda} \in \mathcal{D}, \tag{3.16}$$

where $\mathcal{D}$ is a bounded domain in $\mathbb{R}^n$ with smooth boundary $\partial\mathcal{D}$, $\boldsymbol{D}$ is a positive $m \times m$ diagonal matrix and $\boldsymbol{g} : \mathbb{R}^m \mapsto \mathbb{R}^m$ is locally Lipschitz. Furthermore, $\boldsymbol{B}$ is the diagonal operator

$$B_i w_i = \tau_i w_i + \beta\frac{\partial w_i}{\partial \eta},\ i = 1, 2, \ldots, m, \tag{3.17}$$

where $\partial/\partial\eta$ denotes the derivative with respect to the outward unit normal on $\partial\mathcal{D}$, $\tau = [\tau_1, \ldots, \tau_m] \in \mathbb{R}_+^m$ and $\beta \in \{0, 1\}$. If the conditions

(i) if $\beta = 0$ implies that $\tau_i = 1$ for all $i = 1, 2, \ldots, m$, or if $\tau_i = 0$ for some $i \in \{1, \ldots, m\}$ implies that $\tau \equiv 0$ and $\beta = 1$, and

(ii) $\boldsymbol{w}_0 \in L^\infty(\mathcal{D}, \mathbb{R}^m)$

hold, then local existence and uniqueness of solutions to (3.14)–(3.16) is guaranteed by the following result [103].

Theorem 3.8 *(Existence and uniqueness of solutions to* (3.14)–(3.16) *[103])*
If conditions (i) and (ii) hold, then there exists a real number $T_{\max} > 0$ and an m-component function $\boldsymbol{N} \in C([0, T_{\max}), \mathbb{R}^m)$ such that (6.1)–(6.22) *has a unique, classical, noncontinuable solution $\boldsymbol{w}(\boldsymbol{\lambda}, t)$ on $\mathcal{D} \times [0, T_{\max})$ and $||w_i(\cdot, t)||_\infty \leq N_i(t)$ on $\mathcal{D}$ for all $i = 1, 2, \ldots, m$ and $t \in [0, T_{\max})$. Furthermore, if $T_{\max} < \infty$, then $||w_i(\cdot, t)||_\infty \to \infty$ on $\mathcal{D}$ for some $i \in \{1, \ldots, m\}$.*

3.4 Stability of equilibria

Many of the differential equations considered in this thesis contain a first order derivative with respect to time. An *equilibrium solution* or *steady state solution* to such a differential equation refers to a solution that is constant over time. An equilibrium solution is therefore any solution for which the time derivative is identically zero. An equilibrium solution may be classified as either *stable* or *unstable*. It is *stable* if solutions which are in some sense close to the equilibrium solution remain close to the equilibrium solution for all time [133]. Certain variations in the precise definition are possible, depending on how the notions of closeness and remaining close are defined. A method for defining and determining the stability of equilibrium solutions to certain non-linear PDEs (as outlined in [47]), of which the equations considered in this thesis are special cases, is outlined in this section.

Suppose a non-linear PDE $\partial\boldsymbol{w}/\partial t = F(\boldsymbol{w})$, together with boundary conditions, has an equilibrium solution $\boldsymbol{w}_e$ satisfying $\boldsymbol{F}(\boldsymbol{w}_e) = 0$, and let $\tilde{\boldsymbol{w}}$ represent a perturbation from $\boldsymbol{w}_e$. The evolution of $\tilde{\boldsymbol{w}}$ over time is governed by $\partial\tilde{\boldsymbol{w}}/\partial t = F(\boldsymbol{w}_e + \tilde{\boldsymbol{w}})$, since $\partial\boldsymbol{w}_e/\partial t = 0$. The former equation, together with the boundary conditions, is called the *non-linear stability problem*. This problem may be approximated by a linear problem if it is assumed that the perturbation is sufficiently small. Let $\tilde{\boldsymbol{w}} = \epsilon\boldsymbol{v}$, where ϵ is a positive real number representing the amplitude of the initial perturbation. Then,

$$F(\boldsymbol{w}_e + \epsilon\boldsymbol{v}) = F(\boldsymbol{w}_e) + \epsilon L\boldsymbol{v} + \mathcal{O}(\epsilon),$$

where L is a linear operator which represents $F(\boldsymbol{w})$ evaluated at $\boldsymbol{w} = \boldsymbol{w}_e$ and $\mathcal{O}(\epsilon)$ represents all terms of order smaller than ϵ. If a solution to the linear perturbation equation

$$\frac{\partial\boldsymbol{v}}{\partial t} = L\boldsymbol{v} \tag{3.18}$$

together with the boundary conditions (which may also be linearised if necessary) grows unbounded, then $\boldsymbol{w}_e$ is unstable; otherwise it is stable. If the solutions to (3.18) together with the (possibly linearised) boundary conditions decay to zero over time, then $\boldsymbol{w}_e$ is called *asymptotically stable*.

3.5 Numerical solution techniques

There exist large classes of PDEs for which analytic solution techniques are not known [23]. Furthermore, examples of solutions to PDEs may guide the intuition of a mathematician in the analysis of a given model. For these reasons, standard methods from numerical analysis are often utilised to approximate solutions to initial-boundary value problems involving PDEs. The most commonly used techniques may be categorised into three classes, namely *finite difference methods*, *finite element methods*, and *spectral methods* [85].

The selection of a suitable numerical solution technique for a particular PDE may depend on various criteria. The ease of implementation and computational efficiency of a numerical method, as well as the accuracy of the method, are to be considered. Furthermore, the characteristics of the PDE under consideration may disqualify the use of certain numerical solution techniques.

Spectral methods are generally the most computationally efficient and the most accurate methods, but are constrained to specific problem classes [85]. The best known spectral technique, the *Fourier transform*, may for example only be applied to solve problems with periodic boundary conditions [25]. When applying these techniques, the solution is expressed as the sum of certain basis functions (such as sinusoids in the case of Fourier series). Other spectral techniques typically used in respect of specific problem classes involve the use of Chebyshev polynomials for bounded one-dimensional problems [98], Bessel functions for two-dimensional radial problems [148], or Legendre polynomials for three-dimensional problems involving Laplace's equation[2] [67], to name but a few.

Finite element methods are typically the most taxing from a computational point of view, but may be applied to an extensive class of problems with complicated boundary conditions [140]. These techniques may, for example, be used when the problem domain has an irregular geometry. In these techniques, the domain is discretised into a set of nonoverlapping polyhedrons (such as triangles) [139]. Interpolation functions are then determined on this discretised domain with a view to minimise some error function, as an approximate solution [30].

The focus in this thesis is, however, on finite difference methods, since these techniques are best suited to solving the models considered, for reasons which will become apparent later. These techniques are computationally more efficient than finite element methods and various types of boundary conditions are tractable under these routines [85].

Finite difference methods approximate derivative values over a discretised domain using difference equations. The required equations may be derived from Taylor expansions (see Theorem 3.4). For example, an expression approximating the value of the first derivative of a function may be derived from the Taylor expansions [133]

$$h(\lambda+\delta) = h(\lambda) + \delta h'(\lambda) + \frac{\delta^2}{2!}h''(\lambda) + \frac{\delta^3}{3!}h'''(\lambda+\alpha_1\delta) \tag{3.19}$$

and

$$h(\lambda-\delta) = h(\lambda) - \delta h'(\lambda) + \frac{\delta^2}{2!}h''(\lambda) - \frac{\delta^3}{3!}h'''(\lambda+\alpha_2\delta), \tag{3.20}$$

where $\alpha_1, \alpha_2 \in [0, 1]$. Subtracting (3.20) from (3.19) yields

$$h(\lambda+\delta) - h(\lambda-\delta) = 2\delta h'(\lambda) + \frac{\delta^3}{3!}\left(h'''(\lambda+\alpha_1\delta) + h'''(\lambda+\alpha_2\delta)\right) \tag{3.21}$$

[2] Laplace's equation, defined as $\nabla w = 0$, arises in many applications in science, including electromagnetism, astronomy and fluid dynamics [151].

and it follows from the mean value theorem (see Theorem 3.1) that

$$h'''(\lambda_a) = \frac{h'''(\lambda + \alpha_1\delta) + h'''(\lambda + \alpha_2\delta)}{2}$$

for some $\lambda_a \in [\lambda, \lambda + \delta]$. By substituting this expression into (3.21) and dividing by 2δ, the expression

$$h'(\lambda) = \frac{h(\lambda + \delta) - h(\lambda - \delta)}{2\delta} + \frac{\delta^2}{6} h'''(\lambda_a) \tag{3.22}$$

for the first derivative of h at λ is obtained. The first derivative of h at λ may therefore be approximated by

$$h'(\lambda) \approx \frac{h(\lambda + \delta) - h(\lambda - \delta)}{2\delta} \tag{3.23}$$

with a truncation error of order δ^2. By utilising (3.23) on a discretised domain, the value of the first derivative of a function at a given point is therefore approximated using the neighbouring function values. The accuracy of the approximation may be improved either by utilising more terms in the Taylor expansions during the derivation of an approximation scheme, thus increasing the order of the truncation error, or by decreasing the discretisation step size δ. Examples of such centre-difference schemes are provided in Table 3.1. For approximations at the boundaries, similar expressions may be derived to ultimately utilise only forward or backward differences, as shown in Table 3.2.

Centre difference schemes	Truncation error
$h'(\lambda) \approx {h(\lambda+\delta) - h(\lambda-\delta)}/{2\delta}$	$\mathcal{O}(\delta^2)$
$h''(\lambda) \approx {h(\lambda+\delta) - 2h(\lambda) + h(\lambda-\delta)}/{\delta^2}$	$\mathcal{O}(\delta^2)$
$h'(\lambda) \approx {-h(\lambda+2\delta) + 8h(\lambda+\delta) - 8h(\lambda-\delta) + h(\lambda-2\delta)}/{12\delta}$	$\mathcal{O}(\delta^4)$
$h''(\lambda) \approx {-h(\lambda+2\delta) + 16h(\lambda+\delta) - 30h(\lambda) + 16h(\lambda-\delta) - h(\lambda-2\delta)}/{12\delta^2}$	$\mathcal{O}(\delta^4)$

TABLE 3.1: *Centre-difference schemes for the approximation of first and second derivatives and the order of their truncation errors.*

$\mathcal{O}(\delta^2)$ forward- and backward-difference schemes
$h'(\lambda) \approx {-3h(\lambda) + 4h(\lambda+\delta) - h(\lambda+2\delta)}/{2\delta}$
$h'(\lambda) \approx {3h(\lambda) - 4h(\lambda-\delta) + h(\lambda-2\delta)}/{2\delta}$
$h''(\lambda) \approx {2h(\lambda) - 5h(\lambda+\delta) + 4h(\lambda+2\delta) - h(\lambda+3\delta)}/{\delta^3}$
$h''(\lambda) \approx {2h(\lambda) - 5h(\lambda-\delta) + 4h(\lambda-2\delta) - h(\lambda-3\delta)}/{\delta^3}$

TABLE 3.2: *Forward- and backward-difference schemes for the approximation of first and second derivatives with an $\mathcal{O}(\delta^2)$ truncation error.*

The aforementioned schemes are usually implemented using sparse matrices [85]. The points in the discretised domain are stacked in a vector (if the original domain has more than one dimension, since in the one-dimensional case, the discretised domain is a vector). A sparse matrix which performs derivative approximations on these points upon multiplication may then be constructed. Suppose, as an example, that the second order derivative of a real-valued function w defined on the domain $[\lambda_a, \lambda_b] \subset \mathbb{R}$ is to be approximated, and that periodic Dirichlet boundary conditions are imposed. Multiplying by $(1/\delta^2)\boldsymbol{D_2} \cdot \boldsymbol{w}$ approximates the operation $\mathrm{d}^2/\mathrm{d}\lambda^2$ on the discretised domain $\boldsymbol{x} = [\lambda_a, \lambda_a + \delta, \lambda_a + 2\delta, \ldots, \lambda_b - \delta]$, where $\boldsymbol{w}$ is a discretised approximation of $w(\lambda, t)$ at each point in $\boldsymbol{x}$ at time t for $\delta = \frac{\lambda_b - \lambda_a}{N}$, where N is the number of discretisation points, and

$$\boldsymbol{D_2} = \begin{bmatrix} -2 & 1 & 0 & 0 & \dots & 1 \\ 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \ddots & & 0 \\ 0 & 0 & \ddots & & & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & \dots & 1 & -2 & 1 \\ 1 & 0 & 0 & \dots & 1 & -2 \end{bmatrix} \tag{3.24}$$

is an $N \times N$ matrix. The point λ_b is excluded from $\boldsymbol{x}$ since $w(\lambda_a, t) = w(\lambda_b, t)$ because of the periodicity of w on $[\lambda_a, \lambda_b]$. Slight alterations are made when implementing different boundary conditions. By including λ_b at the end of $\boldsymbol{x}$ (and therefore also increasing the dimension of $\boldsymbol{w}$ by 1), changing D_2 to a $(N+1) \times (N+1)$ square matrix and changing the elements in row 1, column 2 and row $N+1$, column N to 2, for example, zero-flux Neuman boundary conditions are imposed.

Parabolic or hyperbolic PDEs describing physical systems tend to have a first derivative with respect to time [133]. By utilising finite difference approximations of all spatial derivatives, standard time-stepping techniques for ODEs may therefore be utilised to solve initial-boundary value problems involving these types of PDEs. The simplest time-stepping technique is known as *Euler's method* [85]. In this technique, the definition of a derivative $\mathrm{d}w/\mathrm{d}t = \lim_{\triangle t \to 0}(w(t + \triangle t) - w(t))/\triangle t$, is utilised to approximate solutions to ODEs. Consider the ODE

$$\frac{\mathrm{d}w}{\mathrm{d}t} = h(w, t) \tag{3.25}$$

in which case

$$\frac{w(t + \triangle t) - w(t)}{\triangle t} \approx h(w, t). \tag{3.26}$$

By rearranging the terms of (3.26), it follows that

$$w(t + \triangle t) \approx w(t) + \triangle t h(w, t). \tag{3.27}$$

Given an initial condition $w(0) = w_0$, a solution to (3.25) for $t > 0$ may therefore be approximated using (3.27). As with finite difference approximation, time-stepping schemes may be derived from Taylor expansions. The number of terms utilised in these Taylor expansions once again determine the order of the truncation error made when using the scheme. In general, methods used to iterate solutions to differential equations forward in time utilising a single initial point are referred to as *Runga-Kutta* methods [85]. Such methods, which express the solution at the next time step as a function of the current solution, are called *explicit schemes*, in contrast to *implicit* iteration schemes which utilise the future solution on both sides of the difference equation.

A balance between ease of implementation, accuracy and computational efficiency is usually sought in numerical schemes. When considering accuracy, the round-off error made during each time step must be considered in conjunction with the truncation error. A smaller step size decreases the truncation error, but increases the cumulative effect of round-off errors. For a given numerical scheme and a fixed local round-off error size, the global accuracy therefore cannot be increased above a certain threshold. Further increases in accuracy are available by deriving schemes with higher order truncation errors by utilising more terms of the Taylor expansion. Perhaps the most popular explicit time-stepping scheme is the 4^{th}*-order Runga-Kutta method* [30],

$$w_{n+1} = w_n + \frac{\triangle t}{6}(h_1 + 2h_2 + 2h_3 + h_4), \tag{3.28}$$

where

$$\begin{aligned} h_1 &= h(t_n, w_n), \\ h_2 &= h\left(t_n + \frac{\triangle t}{2}, w_n + \frac{\triangle t}{2}h_1\right), \\ h_3 &= h\left(t_n + \frac{\triangle t}{2}, w_n + \frac{\triangle t}{2}h_2\right) \text{ and} \\ h_4 &= h(t_n + \triangle t, w_n + \triangle t h_3). \end{aligned}$$

This scheme has a local truncation error of $\mathcal{O}\triangle t^5$ and therefore a global truncation error of $\mathcal{O}\triangle t^4$. If the local truncation error of a difference equation approximating a solution to a PDE initial and boundary value problem tends to zero as the time and space discretisation intervals tend to zero, then the difference equation is said to be *consistent* [10]. The notion of consistency is formally defined as follows.

Definition 3.12 *(Consistency [133, p. 40–41])*
Suppose w is the exact solution to a PDE together with appropriate boundary and initial conditions, involving two independent variables (e.g. space and time), defined by $L(w) = 0$, and that W is the exact solution to a difference equation defined by $P(W) = 0$. The local truncation error at a point (i, j) is then given by $T_{i,j}(w) = P_{i,j}(w)$. If $T_{i,j}(w) \to 0$ as $\triangle t \to 0$ and $\triangle\lambda \to 0$, where $\triangle t$ and $\triangle\lambda$ represent the time and space discretisation step sizes, respectively, then P is consistent with L.

A numerical scheme is considered *stable* [133] if (truncation and round-off) errors decay in proportion to the exact solution of the PDE, as the scheme progresses forward in time. Implicit schemes, which are typically more cumbersome to implement than explicit schemes, are usually selected for their excellent stability properties [85]. According to the *Lax equivalence theorem* [31], the discretised approximation to a well posed initial-value problem converges to the exact solution of the initial-value problem if it is consistent and stable. The notion of convergence of a difference equation is defined as follows.

Definition 3.13 *(Convergence [133, p. 43–44])*
Suppose w is the exact solution to a PDE together with appropriate boundary and initial conditions, involving two independent variables (e.g. space and time) and defined by $L(w) = 0$, and that W is the exact solution of a difference equation defined by $P(W) = 0$. If at a time $t = t_a$ the solution $W \to w$ as $\triangle t \to 0$ and $\triangle\lambda \to 0$, where $\triangle t$ and $\triangle\lambda$ represent the time and space discretisation step sizes, respectively, then P is convergent with respect to L.

During numerical analyses of equations of the form (3.12) it has been found that explicit schemes are efficient solution methods [142]. It has been shown that a difference scheme for approximating solutions to equations of the form (3.12) based on (3.27) and the $\mathcal{O}(\delta^2)$ centre difference approximation of the second order spatial derivatives is convergent for an abitrary number of spatial dimensions and for any number of solution components, if the exact solution of the PDE initial-boundary value problem has a continuous second order derivative with respect to time, as well as a continuous fourth order derivative with respect to space, for each component [89].

In the simplest case of a single component in one spatial dimension, the scheme is defined as

$$\boldsymbol{w}_{n+1} = \boldsymbol{A}\boldsymbol{w}_n + \triangle t g(\boldsymbol{w}_n), \tag{3.29}$$

where $\boldsymbol{A}$ is a difference operator, $\boldsymbol{w}$ is a vector approximation of the component densities along the domain at a given point in time and $\triangle t$ is the time discretisation step size. In this one-dimensional case, the difference operator is defined as

$$\boldsymbol{A} = \boldsymbol{I} + \nu \boldsymbol{D_2} = \begin{bmatrix} 1-2\nu & \nu & 0 & 0 & \dots & 0 \\ \nu & 1-2\nu & \nu & 0 & \dots & 0 \\ 0 & \nu & \ddots & \ddots & & 0 \\ 0 & 0 & \ddots & & & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & \dots & \nu & 1-2\nu & \nu \\ 0 & 0 & 0 & \dots & \nu & 1-2\nu \end{bmatrix},$$

where $\boldsymbol{I}$ is the identity matrix, $\nu = (d\triangle t)/\triangle\lambda^2$, d is the diffusion rate and $\triangle\lambda$ is the spatial discretisation interval size. It has also been shown that this scheme is stable if $\nu \leq 1/2$ [89].

3.6 Chapter summary

A variety of basic mathematical definitions and theorems used in the arguments put forward later in this thesis were reviewed in the first section of this chapter. A brief introduction to the field of differential equations was then given, after which a standard classification of second-order linear PDEs was reviewed. The focus of the discussion next shifted to reaction-diffusion systems. A method for analysing the stability of equilibrium solutions to certain PDEs was then reviewed. Certain numerical solution techniques for the approximation of solutions to initial-boundary value problems involving PDEs were then described briefly, after which finite difference methods were reviewed.

The mathematical literature review in this chapter is by no means exhaustive or in any sense representative. Instead, the contents of this chapter should be understood to form a prerequisite basis of understanding for the remainder of the thesis, which will contain frequent references back to this chapter.

Part II

Mathematical investigation

CHAPTER 4

Models of wealth distribution

Contents

Two simple, highly idealised models of the dynamics of wealth distribution with contrasting redistribution schemes are derived in this chapter. The first is a reaction-diffusion model representing a trickle-down type of wealth redistribution, while the second model represents the imposition of a linear tax rate over wealth. First, the assumptions made in order to derive the mathematical models are noted, after which the mathematical derivation follows. Metrics that may be used to characterise the growth of wealth and the extent of inequality in model solutions are defined next, and their equivalence to traditional measures of wealth distribution is demonstrated. The wealth growth functions to be considered are then introduced. The existence and uniqueness of model solutions, as well as the nonnegativity of these solutions and their first derivatives over space, are established next. The models' limitations and intended use are finally discussed in a formal model apology.

4.1 Model assumptions

The following assumptions are made in order to derive two models of the temporal evolution of wealth distribution in a socio-economic system:

1. *The notion of space.* It is assumed that the distribution of wealth in an economy can be represented over some hypothetical space representing equally distributed entities (which may be companies within a certain sector, persons or households, for example). A certain portion of this space therefore refers to that same portion of the total entities.

2. *Continuity and smoothness of wealth distribution over space and time.* It is assumed that a sufficiently large number of entities may be ordered in such a manner that their wealth can be approximated by a twice differentiable continuous function over space and time. Closeness of entities in space therefore implies similar wealth positions.

3. *Conservation of wealth.* It is assumed that wealth is conserved in redistributive processes, so that the change in total wealth depends explicitly on growth processes only. Since the wealth distributional shape may influence the growth rates attained, redistribution does implicitly influence the change in wealth over time, but wealth is not created or lost during transfers.

4. *Homogeneity in the growth and redistribution of wealth.* As the work in this thesis is an exploration of the simplest possible environments facilitating the emergence of inequality, it is assumed that wealth growth characteristics depend only on the current wealth distribution. The possibility of total wealth equality is allowed in the exploration of inequality, and it follows that growth and redistribution processes affecting an entity do not explicitly depend on the entity's position in space.

5. *Initial wealth distribution.* Since entities may be ordered according to their wealth, it is assumed without loss of generality that their initial wealth distribution is non-decreasing over space. It is also assumed that this initial distribution is non-negative, and so entities may posses zero, but not negative, wealth.

6. *Redistribution of wealth.* It is assumed that, through regulation, some wealth may be transferred from the wealthy to the less wealthy. Two scenarios are considered:

   (a) It is assumed that the rate of wealth redistribution is directly proportional to the wealth distribution gradient, with a constant of proportionality $d \geq 0$. Since wealth is represented by a continuous, twice differentiable function over space, this translates to a subtraction from the wealth of entities at and near local wealth maxima and an addition to the wealth of entities at or near local wealth minima, separated by points of inflection. As mentioned in §2.3, this approach has previously been adopted in modelling wealth redistribution. This is mathematically equivalent to the process of diffusion, and the net redistribution effect therefore corresponds to wealth only flowing directly between entities in close proximity (*i.e.* having similar wealth positions). It should be noted that actual transfers need not occur in this fashion, but that the positions of progressive tax brackets and of grant recipients may be determined using the second derivative of an estimate of the current wealth distribution over space, resulting in a diffusion-like net redistributive effect. The model is in this case referred to as the *trickle-down model*.

   (b) It is assumed that the rate of redistribution experienced by an entity is directly proportional to the entity's wealth position relative to the population's average wealth position, also with a constant of proportionality $d \geq 0$. This corresponds to a non-localised wealth redistribution scheme where the greatest net contribution, and net benefit, occur at the most wealthy and least wealthy entity, respectively. This corresponds to a linear tax over wealth, with negative taxes at wealth levels below the mean wealth, so that redistribution is conservative as required by Assumption 3. The model in this case is referred to as the *linear redistribution model*

7. *Boundary conditions.* In the space used to represent entities in the system, wealth flux across the boundaries would represent unaccounted for changes in total wealth, as a result of changes in the wealth positions of only the most and least wealthy entities in

the population. This would translate into a non-wealth-conserving redistribution scheme. Zero-flux Neuman boundary conditions are therefore imposed for case (a) above, meaning that wealth cannot exit or enter the system across its boundaries. In case (b), however, the redistribution term implicitly conserves wealth, so that no further restrictions on the boundaries are required.

8. *The growth of wealth.* It is assumed that, viewed from a macro perspective, the growth experienced at a given wealth level is proportional to that wealth level. This does not necessarily imply that individual entities with the same wealth level experience the same growth (entities' positions in space are not necessarily fixed), but that distributional changes over time can be captured by describing the average per capita wealth growth rate in the wealth distribution at a given wealth attainment level. The growth function therefore assumes Kolmogorov form [42]. It is assumed that the distributional dependence of the per capita growth rate may be captured by adopting a measure of the distributional shape (the normalised variance is chosen in this case), and of the total wealth.

4.2 Model derivation

Two mathematical models of wealth dynamics are derived in this section, based on the assumptions outlined above. First, a model with a trickle-down type redistribution, inspired by the notion of diffusion, is derived. An alternative model, employing a conservative linear redistributive tax scheme, is then derived.

4.2.1 Trickle-down model

By assumptions 1 and 2 of §4.1, the wealth level at a position λ in a finite, one-dimensional real spatial domain $\mathcal{D}$ and at a time $t \in [0, \infty)$ may be denoted by a twice differentiable continuous function $w(\lambda, t)$. If $\mathcal{V} \subset \mathcal{D}$ is an arbitrary subset of $\mathcal{D}$, with boundaries λ_a and λ_b, where $\lambda_a < \lambda_b$, and the net change in wealth as a result of redistribution is equivalent to transfers performed only between adjacent entities according to the first case of Assumption 6(a), then

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{V}} w(\lambda, t)\,\mathrm{d}v = J(\lambda_a, t) - J(\lambda_b, t) + \int_{\mathcal{V}} g(w, \mu, \overline{V})\,\mathrm{d}v,$$

according to Assumption 3, where $J(\lambda, t)$ is the flux of wealth at position λ and time t, and g is the net growth in $\mathcal{V}$, with μ and $\overline{V}$ representing the total wealth and the normalised variance of w over $\mathcal{D}$, respectively. According to Assumption 6, $J(\lambda, t) = -d\partial w/\partial\lambda$, where d is a constant of proportionality denoting the diffusion rate, or the rate of redistribution of wealth. Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{V}} w(\lambda, t)\,\mathrm{d}v = -d \left[\left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=\lambda_a} - \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=\lambda_b} \right] + \int_{\mathcal{V}} g(w, \mu, \overline{V})\,\mathrm{d}v$$

and so by the fundamental theorem of the calculus (see Theorem 3.2),

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{V}} w(\lambda, t)\,\mathrm{d}v = d \int_{\mathcal{V}} \frac{\partial^2 w}{\partial \lambda^2}\,\mathrm{d}v + \int_{\mathcal{V}} g(w, \mu, \overline{V})\,\mathrm{d}v,$$

or, in rearranged form,

$$\int_{\mathcal{V}} \left[\frac{\partial w}{\partial t} - g(w, \mu, \overline{V}) - d\frac{\partial^2 w}{\partial \lambda^2} \right] \mathrm{d}v = 0.$$

Since $\mathcal{V}$ is an arbitrary subset of $\mathcal{D}$, however, it follows that

$$\frac{\partial w}{\partial t} = g(w, \mu, \overline{V}) + d\frac{\partial^2 w}{\partial \lambda^2}.$$

Furthermore, since the growth rate is expressed as a function of the local wealth attainment level w, the total wealth μ and the normalised variance $\overline{V}$, the reaction term $g(w, \mu, \overline{V})$ may be written in Kolmogorov form $wf(w, \mu, \overline{V})$ according to Assumption 8. Assume, without loss of generality, that $\mathcal{D}$ is the unit interval $[0, 1]$ of real numbers. Then the boundary value problem

$$\text{D.E.} \quad \frac{\partial w}{\partial t} = wf(w, \mu, \overline{V}) + d\frac{\partial^2 w}{\partial \lambda^2} \quad \lambda \in \mathcal{D},\ t \geq 0, \tag{4.1}$$

$$\text{B.C.} \quad \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=0} = \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=1} = 0,\ t \geq 0, \tag{4.2}$$

$$\text{I.C.} \quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0, \quad \lambda \in \mathcal{D}, \tag{4.3}$$

is obtained as a model of wealth distribution over time, according to Assumptions 5 and 7, where f is the per capita growth rate of wealth and w_0 is the initial wealth distribution. The model (4.1)–(4.3) is henceforth called the *trickle-down model* of wealth distribution.

4.2.2 Linear redistribution model

In the second case of Assumption 6 of §4.1, it follows from Assumptions 1 and 2 that the wealth level at a position λ in a finite, one-dimensional spatial domain $\mathcal{D} \equiv [0, 1]$ and at a time $t \in [0, \infty)$ may once again be denoted by a twice differentiable continuous function $w(\lambda, t)$. If $\mathcal{V} \subset \mathcal{D}$ is again an arbitrary subset of $\mathcal{D}$, and wealth is transferred between entities depending on their position relative to the population's mean wealth according to Assumption 6(b), then

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{V}} w(\lambda, t)\,\mathrm{d}v = d\int_{\mathcal{V}} [\mu(t) - w(\lambda, t)]\,\mathrm{d}v + \int_{\mathcal{V}} g(w, \mu, \overline{V})\,\mathrm{d}v,$$

according to Assumption 3, where $\mu(t) = \int_{\mathcal{D}} w(\lambda, t)\,\mathrm{d}\lambda$, which may be rearranged to obtain

$$\int_{\mathcal{V}} \left[\frac{\partial w}{\partial t} - g(w, \mu, \overline{V}) - d[\mu(t) - w(\lambda, t)]\right]\mathrm{d}v = 0.$$

Since $\mathcal{V}$ is an arbitrary subset of $\mathcal{D}$, it follows that

$$\frac{\partial w}{\partial t} = g(w, \mu, \overline{V}) + d[\mu(t) - w(\lambda, t)].$$

Furthermore, since the effects of growth are expressed as a growth rate in terms of a given wealth attainment, the reaction term $g(w, \mu, \overline{V})$ may again be written in Kolmogorov form $wf(w, \mu, \overline{V})$ according to Assumption 8, in which case the initial value problem

$$\text{D.E.} \quad \frac{\partial w}{\partial t} = wf(w, \mu, \overline{V}) + d[\mu(t) - w(\lambda, t)] \quad \lambda \in \mathcal{D},\ t \geq 0, \tag{4.4}$$

$$\text{I.C.} \quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0, \quad \lambda \in \mathcal{D} \tag{4.5}$$

is obtained as a model of wealth distribution over time, according to Assumptions 5 and 7, where f is the per capita growth rate of wealth and w_0 is the initial wealth distribution. The model (4.4)–(4.5) is henceforth called the *linear redistribution model* of wealth distribution.

4.3 Measures of the distribution of wealth

At an arbitrary time $t \geq 0$, the total wealth is given by

$$\mu(t) = \int_{\mathcal{D}} w(\lambda, t)\, \mathrm{d}\lambda. \tag{4.6}$$

This wealth (which is also the mean wealth because of the choice of $\mathcal{D}$) is analogous to

$$\mu(t) = \int_0^\infty w(\lambda, t)\theta(w, t)\, \mathrm{d}w$$

by the following proposition, where θ is the density function of w at time t.

Proposition 4.1 *(Equivalence of parade and frequency distribution modelling approaches) Let $w(\lambda, t)$ be a solution to (4.1)–(4.3) or (4.4)–(4.5) on $\mathcal{D} \times [0, T_{\max}]$ and suppose $\theta(w, t)$ is the density function of w at time $t \in [0, T_{\max}]$ for some $T_{\max} > 0$. Then*

$$\int_{\mathcal{D}} w^n(\lambda, t)\, \mathrm{d}\lambda = \int_0^\infty w^n \theta(w, t)\, \mathrm{d}w \tag{4.7}$$

for all $t \in [0, T_{\max}]$ and any $n \in \mathbb{N}$.

Proof. Viewing the integral on the left-hand side of (4.7) as a Riemann sum[1] it follows that

$$\int_{\mathcal{D}} w^n(\lambda, t)\, \mathrm{d}\lambda = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} w^n(\lambda_i, t), \tag{4.8}$$

where $\lambda_i = i/m$ for $i = 0, 1, \ldots, m$, as illustrated in Figure 4.1.

Denote the minimum and maximum value in the set $\{w^n(\lambda_i, t)\}_{i=1}^m$ by $w^n_{\min}(t)$ and $w^n_{\max}(t)$, respectively, and partition the interval $[w^n_{\min}(t), w^n_{\max}(t)]$ into p wealth brackets of equal length. Let $w^n_j(t)$ be the centre point of the j-th wealth bracket and denote the number of $w^n(\lambda_i, t)$ terms on the right-hand side of (4.8) within the j-th bracket by $\gamma_j(t)$. Then

$$\begin{aligned}
\int_{\mathcal{D}} w^n(\lambda, t)\, \mathrm{d}\lambda &= \lim_{m\to\infty} \lim_{p\to\infty} \frac{w_{\max}(t) - w_{\min}(t)}{mp} \sum_{j=1}^{p} \gamma_j(t) w^n_j(t) \\
&= \lim_{m\to\infty} \lim_{p\to\infty} \frac{w_{\max}(t) - w_{\min}(t)}{p} \sum_{j=1}^{p} \theta_j(t) w^n_j(t) \\
&= \int_0^\infty w^n \theta(w)\, \mathrm{d}w,
\end{aligned}$$

where $\theta_j(t) = \gamma_j(t)/m$ denotes the proportion of $w^n(\lambda_i, t)$ terms on the right-hand side of (4.8) that reside within the j-th wealth bracket at time t, and $\theta(w, t)$ is the density function of w at time t. □

[1] The sum $\sum_{i=1}^{n} h(\lambda_i^*)\triangle\lambda_i$ of values of a real-valued function h defined on a real, closed interval $[\lambda_a, \lambda_b]$ partitioned by points $\lambda_a < \lambda_1 < \lambda_2 < \cdots < \lambda_{n-1} < \lambda_b$ into n intervals of lengths $\triangle\lambda_1, \triangle\lambda_2, \ldots, \triangle\lambda_n$, where λ_k^* is any point in subinterval $k \in \{1, \ldots, n\}$, is called a *Riemann sum* [136].

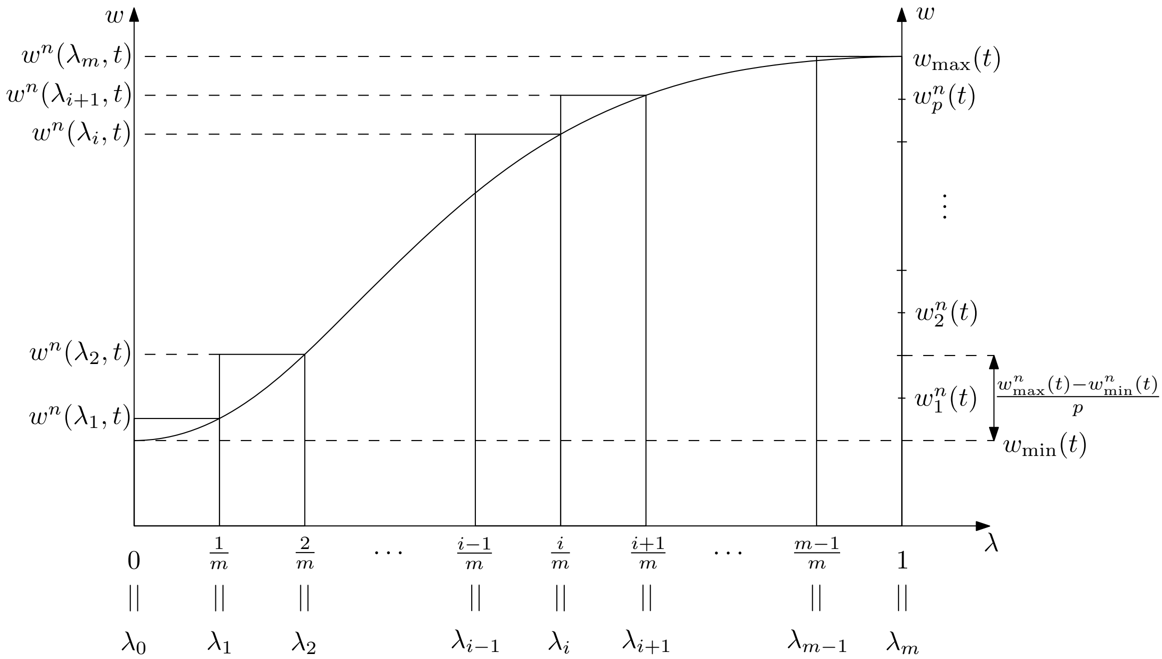


FIGURE 4.1: *Viewing an integral as a Riemann sum.*

Consider the variance

$$V(t) = \int_{\mathcal{D}} [w(\lambda, t) - \mu(t)]^2 \,\mathrm{d}\lambda \tag{4.9}$$

$$= \int_{\mathcal{D}} w^2(\lambda, t) \,\mathrm{d}\lambda - 2\mu(t) \int_{\mathcal{D}} w(\lambda, t) \,\mathrm{d}\lambda + \mu^2(t) \tag{4.10}$$

of the wealth function $w(\lambda, t)$ over $\mathcal{D}$ at a fixed value of t. By substituting (4.6) into (4.10), it follows that

$$V(t) = \int_{\mathcal{D}} w^2(\lambda, t) \,\mathrm{d}\lambda - \mu^2(t), \tag{4.11}$$

and by substituting (4.7) into (4.11), the variance may be rewritten as

$$\begin{aligned} V(t) &= \int_0^\infty w^2\theta(w) \,\mathrm{d}w - \mu^2(t) \\ &= \int_0^\infty [w - \mu(t)]^2\theta(w) \,\mathrm{d}w, \end{aligned}$$

which is a measure of the *absolute inequality* of the wealth function $w(\lambda, t)$ at time t. Absolute inequality of wealth measures the numerical difference between entities' wealth positions. A proportional increase in wealth across the entire domain (multiplication by some positive constant) therefore causes an increase in absolute inequality of wealth, but this absolute inequality is unaffected by a uniform addition of wealth across the domain.

In contrast, the normalised variance

$$\overline{V}(t) = \int_{\mathcal{D}} \left[\frac{w(\lambda, t) - \mu(t)}{\mu(t)}\right]^2 \mathrm{d}\lambda \tag{4.12}$$

is a measure of the *relative inequality* of the wealth function $w(\lambda, t)$ at time t. Here the term *relative inequality* refers to changes in the normalised wealth distribution, and is therefore unaffected by proportional changes in wealth across the entire domain. A uniform addition to the wealth position of all entities will, however, result in a decrease in relative inequality.

It can be shown that (4.12) is a special case of the class of generalised entropy measures of inequality defined in (2.1). This may be confirmed by the expansion

$$\begin{aligned}
\overline{V}(t) &= \int_{\mathcal{D}} \left[\frac{w^2(\lambda,t) - 2w(\lambda,t)\mu(t) + \mu^2(t)}{\mu^2(t)}\right] \mathrm{d}\lambda \\
&= \int_{\mathcal{D}} \left[\frac{w(\lambda,t)}{\mu(t)}\right]^2 \mathrm{d}\lambda - \frac{2}{\mu(t)} \int_{\mathcal{D}} w(\lambda,t)\,\mathrm{d}\lambda + 1 \\
&= \int_{\mathcal{D}} \left[\frac{w(\lambda,t)}{\mu(t)}\right]^2 \mathrm{d}\lambda - 1,
\end{aligned} \tag{4.13}$$

of the expression (4.12). Note that (4.13) is simply the continuous case of (2.1) multiplied by 2, with $s = 2$.

It should be noted that relative inequality of wealth is of far greater significance in the current investigation than absolute inequality of wealth. If general statements are to be made about the shape of a wealth distribution, these have to be phrased in terms of relative inequality, since a wealth distribution of some shape may correspond to an arbitrary absolute inequality of wealth.

While certainly applicable in terms of relative inequality of wealth, the argument that increasing inequality is unsustainable does not necessarily apply to absolute inequality of wealth. For increasing relative inequality of wealth, the relative distribution becomes more unequal over time, meaning that some subset of the population owns an ever-decreasing proportion of the total wealth. It is self-evident that this phenomenon is unsustainable. In a system where net growth is positive and proportional to the current distribution with some constant growth rate, however, the relative wealth distribution remains unchanged, while absolute inequality of wealth is increasing, and it is not evident that this represents an unsustainable situation.

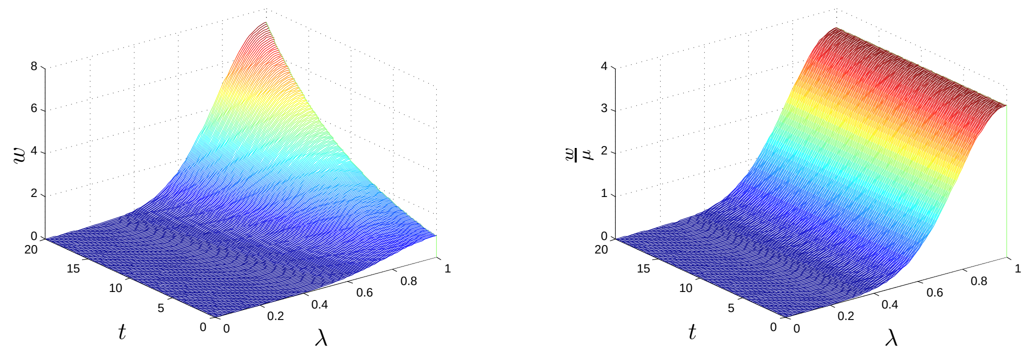
4.4 Per capita wealth growth-rate functions

The different per capita wealth growth-rate functions to be considered in this thesis are now put forward and their underlying assumptions are considered. As the models derived in §4.2 are aimed at facilitating an exploration of the simplest possible descriptions of certain macroeconomic phenomena related to the distribution of wealth, the wealth growth functions to be considered are introduced here and also considered in later chapters in order of increasing complexity. Each new growth function contributes more interesting, and perhaps more realistic, characteristics and behaviour.

As mentioned in §1.1, assuming per capita wealth growth-rate functions that are increasing over wealth is sufficient to produce increasing relative inequality over time. The simplest possible per capita wealth growth-rate function that is increasing over wealth is the linear function

$$f(w) = a + bw, \tag{4.14}$$

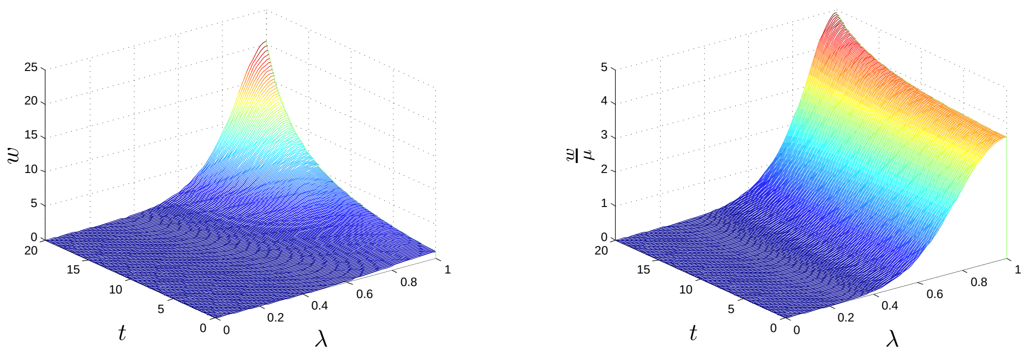
where a and b are constants. If $b = 0$, this represents a case where the entire population experiences the same proportional growth. Absolute inequality increases under this growth function (assuming $a > 0$), since the absolute difference between entities' wealth positions are magnified over time, as illustrated for an initial Gaussian distribution of wealth in Figure 4.2(a). The wealth distribution shape, and its relative inequality, is not influenced by this type of growth, as illustrated in Figure 4.2(b). Any amount of redistribution is therefore expected to result in decreased relative inequality.



(a) *Absolotue wealth distribution over time* (b) *Relative wealth distribution over time*

FIGURE 4.2: *A constant per capita wealth growth rate magnifies the absolute wealth distribution (a), but does not produce a change in the relative wealth distribution (b). The figures are based on numerical approximations of the solution to* (4.1)–(4.3) *obtained by applying the method* (3.29) *with step sizes* $\triangle t = 0.05$ *and* $\triangle\lambda = 1/80$*, and assuming the initial condition* $w_0(\lambda) = \exp(-10(\lambda-1)^2)$ *and parameter values* $a = 0.1$ *and* $b = d = 0$*.*

If $b > 0$, however, the per capita wealth growth-rate function ensures that growth rates at higher wealth levels are greater than at lower wealth levels and therefore an increase in relative inequality is possible, as illustrated for the same Gaussian initial distribution of wealth in Figure 4.3. The extent of redistribution required to limit the growth of relative inequality now becomes a question of interest. The case where $b < 0$ is not of interest in an economic context. Such models (with $a > 0$ and $b < 0$) take the form of Fisher's equation [57] (see equation (3.13)) and have been studied extensively in the literature on biological species modelling. In all subsequent discussions, it is assumed that $b \geq 0$.



(a) *Absolute wealth distribution over time* (b) *Relative wealth distribution over time*

FIGURE 4.3: *A linearly increasing per capita wealth growth rate changes the relative wealth distribution's shape, corresponding to increased inequality. The figures are based on numerical approximations of the solution to* (3.14)–(3.16) *obtained by applying the method* (3.29) *with step size* $\triangle t =$ *and* $\triangle\lambda = 1/80$*, and assuming the initial condition* $w_0(\lambda) = \exp(-10(\lambda-1)^2)$ *and parameter values* $a = 0.1$, $b = 0.01$ *and* $d = 0$*.*

If a is negative, the situation is encountered where negative wealth growth rates are experienced by entities possessing less than a certain critical wealth level. This may represent a situation where there exists a certain minimum capital requirement, below which an entity's fixed expenses are greater than its possible returns, resulting in a steadily decreasing wealth position below this critical wealth level. A question of interest in this situation is, therefore, for a population with a wealth distribution around this critical wealth level, how redistribution impacts the population's capability to rise above this critical wealth level.

An unfortunate property of (4.14), however, is that as entities' wealth levels increase, the rate of growth is allowed to increase indefinitely. Only the initial evolution of inequality subject to this wealth growth-rate function is therefore of interest, as well as changes in the total wealth when distributed around a critical wealth level, whereas continual growth above such a critical level is not of practical interest.

It may be argued that an increase in the total population wealth should result in a corresponding decrease in the growth rate attained at a given fixed wealth level. The emergence of large corporations, which draw greater benefit from economies of scale, detract from the performance of smaller competing companies, for example. A simple addition to (4.14) incorporates this phenomenon,

$$f(w, \mu) = a + bw - c\mu, \tag{4.15}$$

where c is a nonnegative constant and where μ is defined as in (4.6). The per capita wealth growth-rate function (4.15) is therefore capable of capturing, in a very simple fashion, both the instability of an unregulated free market, in the sense that it can allow increasing inequality in the distribution of wealth, and the inherent competition in such a system, where relative value or size is of far greater significance than absolute attainment levels. For both (4.14) and (4.15), a higher level of inequality results in a larger rate of growth of total wealth, for any given mean wealth. While it was noted in §2.2 that increased inequality can lead to increased economic growth, it was also noted that several sources argue that severe inequality of wealth may hamper economic growth. It may therefore be reasonable to assume that inequality contributes to economic growth provided that the level of inequality is less than some threshold, above which increased inequality reduces economic growth. Since the aforementioned wealth growth-rate functions already incorporate the positive relation between increased inequality and growth, only a term which introduces the negative effect of inequality on total growth need be added in order to capture the effect of severe wealth inequality hampering economic growth. A quantitative measure of the relative inequality of a distribution's shape is therefore required.

As was noted in §4.3, the normalised variance corresponds to a *generalised entropy* inequality index (which quantifies relative inequality) with sensitivity parameter $s = 2$. This choice of s weighs wealth deviations farther from the mean more heavily, and results in a metric that is convenient in view of its simplicity. Inherent to this choice is the assumption that a large deviation from the mean wealth in a small population subset is more detrimental to economic growth than a small deviation from the mean in a large population subset, which is reasonable. No claim is made, however, that this choice of sensitivity parameter and corresponding measure of inequality is necessarily superior to other possibilities. The aim is simply to investigate and showcase the variety of possible model solution behaviours that result from the inclusion of a wealth growth function which depends on the extent of wealth inequality. The resulting per capita wealth growth-rate function is

$$f(w, \mu, \overline{V}) = a + bw - c\mu - k\overline{V}, \tag{4.16}$$

where k is a nonnegative constant and where $\overline{V}$ is defined as in (4.12).

The per capita growth-rate functions considered in this thesis are therefore all special cases of (4.16). Figure 4.4 depicts f for positive wealth with $b, c, k > 0$ and $a < c\mu + k\overline{V}$.

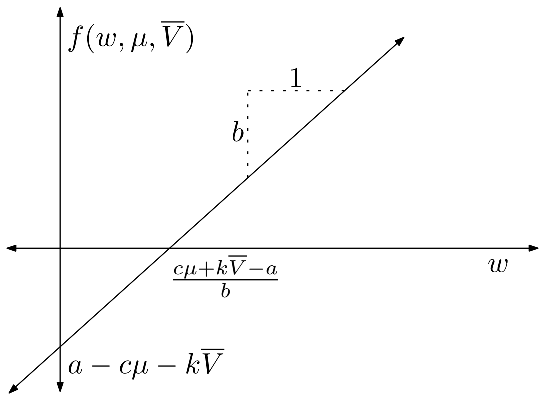


FIGURE 4.4: *Linearly increasing per capita wealth growth-rate as a function of wealth, translating proportional to the mean wealth and to normalised variance.*

4.5 Existence and properties of model solutions

In this section, the existence and uniqueness of solutions to the models derived in §4.2 are established. Nonnegativity of solutions over space is also established in each case.

The system (4.1)–(4.3) is a single-component instance of (3.14)–(3.16), where $\tau \equiv 0$. The following corollary therefore follows from Proposition 3.8.

Corollary 4.1 *(Existence and uniqueness of trickle-down model solutions)*
If w_0 is twice differentiable on $\mathcal{D}$, and $g = wf(w, \mu, \overline{V})$ is Lipschitz continuous on $\mathcal{D} \times [0, T_{\max}]$ for some $T_{\max} > 0$, then (4.1)–(4.3) has a unique, classical solution on $\mathcal{D} \times [0, T_{\max}]$.

The initial value problem (4.4)–(4.5), is a limiting case as $n \to \infty$ of the initial value problem

$$\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t} = \boldsymbol{w} f(\boldsymbol{w}) + d\left[\frac{1}{n}\sum_{i=1}^{n} w_i - \boldsymbol{w}\right], \tag{4.17}$$

$$\boldsymbol{w}(0) = \kappa_i, \quad i \in 1, 2, \ldots n \tag{4.18}$$

for a system of ODEs, where $\boldsymbol{w}$ is an n-column vector, w_i denotes the i-th entry of $\boldsymbol{w}$ and $\kappa_i = w_i(0) = w(i/n, 0)$. The change in height over time of each rectangle in the Riemann sum in (4.17) is determined by an ODE. As $n \to \infty$, the width of each ODE's representative strip tends to zero and the solution to the system (4.17)–(4.18) tends to that of (4.4)–(4.5). According to Proposition 3.6, (4.17)–(4.18) has a unique solution on $[0, T_{\max})$ for some $T_{\max} > 0$ if the right-hand side of (4.17) and its first order derivative over wealth are continuous on $[0, T_{\max})$. This may be extended to the initial value problem (4.4)–(4.5) according to the following corollary.

Corollary 4.2 *(Existence and uniqueness of linear redistribution model solutions)*
The initial value problem (4.4)–(4.5) has a unique solution on $\mathcal{D} \times [0, T_{\max})$ for some $T_{\max} > 0$ if the Riemann sum $\frac{1}{n}\sum_{i=1}^{n} w_i$ in (4.17) converges to $\int_{\mathcal{D}} w \,\mathrm{d}\lambda$. The existence of bounded solutions to (4.17)–(4.18) on $[0, T_{\max})$, according to Proposition 3.7, guarantees the convergence of $\frac{1}{n}\sum_{i=1}^{n} w_i$ on $\mathcal{D} \times [0, T_{\max})$ and therefore the existence and uniqueness of solutions to (4.4)–(4.5) on $\mathcal{D} \times [0, T_{\max})$.

The differential equations (4.1) and (4.4) are both of the form

$$\frac{\partial w}{\partial t} = g(w, \mu, \overline{V}) + dR(w, \mu), \tag{4.19}$$

where $R(w, \mu)$ represents the redistributive influence. The following property of R will be useful in the chapters that follow.

Lemma 4.1 *(Nonnegativity of the redistributive term)*
Let w be the unique solution to the initial-boundary value problem (4.1)–(4.3)*, or to the initial value problem* (4.4)–(4.5)*. Suppose $w(\lambda, t) \geq 0$ for all $\lambda \in \mathcal{D}$ at time t. If $w(\lambda^*, t) = 0$, then* $R|_{w=w(\lambda^*,t)} \geq 0$.

Proof. In the case of (4.1) it follows from the continuity and twice differentiability of w that $\partial^2 w/\partial\lambda^2|_{\lambda=\lambda^*} \geq 0$ by Theorem 3.3, since $w(\lambda^*, t)$ is a minimum of w on $\mathcal{D}$ at time t. Since $g(w, \mu, \overline{V}) = 0$ when $w = 0$, it follows from (4.19) that $\partial w/\partial t|_{\lambda=\lambda^*} \geq 0$. Furthermore, since the minimum value of a function is not greater than its mean, the desired result follows immediately in the case of (4.4). □

The following result now follows from the lemma above.

Proposition 4.2 *(Nonnegativily of model solutions)*
Let $w(\lambda, t)$ be the unique solution to (4.1)–(4.3) *or* (4.4)–(4.5) *on $\mathcal{D} \times [0, T_{\max}]$ for some $T_{\max} > 0$. Then $w(\lambda, t) \geq 0$ on $\mathcal{D} \times [0, T_{\max}]$.*

Proof. By contradiction. Suppose the last time at which w is entirely nonnegative on $\mathcal{D}$, before becoming negative for the first time at some point(s) in $\mathcal{D}$, occurs at $t = t^* \geq 0$. The continuity of w then implies that $w(\lambda^*, t^*) = 0$ and $\partial w/\partial t|_{\lambda=\lambda^*,t=t^*} < 0$ for some $\lambda^* \in \mathcal{D}$. The nonnegativity of the redistribution term $R(w)$ at (λ^*, t^*), according to Lemma 4.1, together with the fact that $g|_{w=w(\lambda^*,t^*)} = 0$ since $g(w, \mu, \overline{V}) = wf(w, \mu, \overline{V})$, however, imply that $\partial w/\partial t|_{\lambda=\lambda^*,t=t^*} \geq 0$, a contradiction. □

It is finally shown that the derivative of wealth over space also remains nonnegative for all time, and therefore that the maximum wealth in (4.19) always occurs at $\lambda = 1$ and that the minimum wealth occurs at $\lambda = 0$.

Proposition 4.3 *(Preservation of wealth ordering over space)*
Let $w(\lambda, t)$ be the unique solution to (4.1)–(4.3) *or* (4.4)–(4.5) *on $\mathcal{D} \times [0, T_{\max}]$ for some $T_{\max} > 0$. Then $\frac{\partial w}{\partial \lambda} \geq 0$ on $\mathcal{D} \times [0, T_{\max}]$.*

Proof. Suppose $\partial w/\partial\lambda \geq 0$ at an arbitrary time $t_a \geq 0$. Then, for arbitrary values $\lambda_a, \lambda_b \in \mathcal{D}$ satisfying $\lambda_a < \lambda_b$, it follows that $w(\lambda_a, t_a) \leq w(\lambda_b, t_a)$. The difference between the rate of change in w at λ_b and λ_a at time t is given by

$$\begin{aligned}\rho(w, \lambda_a, \lambda_b, t) &= \left.\frac{\partial w}{\partial t}\right|_{\lambda=\lambda_b} - \left.\frac{\partial w}{\partial t}\right|_{\lambda=\lambda_a} \\ &= \underbrace{wf(w, \mu, \overline{V})|_{\lambda=\lambda_b} - wf(w, \mu, \overline{V})|_{\lambda=\lambda_a}}_{(*)} + \underbrace{d\left[R(w, \mu)|_{\lambda=\lambda_b} - R(w, \mu)|_{\lambda=\lambda_a}\right]}_{(**)}.\end{aligned}$$

If $w(\lambda_a, t_a) = w(\lambda_b, t_a)$, then it follows that $\partial w/\partial\lambda = 0$ everywhere on $[\lambda_a, \lambda_b]$ and the term $(*)$ equals zero. Since $[\mu(t_a) - w(\lambda_b, t_a)] - [\mu(t_a) - w(\lambda_a, t_a)] = 0$ if $w(\lambda_a, t_a) = w(\lambda_b, t_a)$ it follows that the term $(**)$ is zero in the case of (4.4). In the case of (4.1), furthermore, $\partial^2 w/\partial\lambda^2|_{\lambda=\lambda_a,t=t_a} \leq 0$ and $\partial^2 w/\partial\lambda^2|_{\lambda=\lambda_b,t=t_a} \geq 0$ if $w(\lambda_a, t_a) = w(\lambda_b, t_a)$, from which it follows that $\rho(w, \lambda_a, \lambda_b, t_a) \geq 0$. The nonnegativity of the spatial derivative as a function of λ therefore cannot be reversed. □

4.6 Model apology

The models derived in §4.2 are highly idealised abstractions of the growth and redistribution of wealth in real socio-economic environments for a number of reasons. The use of a continuous function to represent the characteristics of some discrete set of entities is a case in point. Certain requirements in continuous space do not necessarily have interpretations in a discrete paradigm. For example, the requirement of zero-flux Neumann boundary conditions does not have an immediate interpretation in a discrete sense, since the proportion of the wealth function for which the spatial derivative is zero may be arbitrarily small. This requirement does not therefore correspond to an expectation in terms of the shape of wealth distributions observed in (discrete) reality.

Considering a closed system is also idealistic, since it is rarely possible to isolate self-contained economic systems completely. This model aspect may, however, be considered justified in this case, since it represents the simplest possible system, and is commonly used in theoretical models of economic systems [6, 19, 75].

The choice of the notion of space in the models of this chapter corresponds to what is known as the *parade approach* [39], as opposed to working with frequency distributions, which translate into probability density distributions in the continuous case. The parade approach corresponds to the thought experiment of parading a population in a long single-file line arranged according to nondecreasing wealth, with each person's height representing his or her wealth. This approach is intuitively more appealing than the alternative of frequency space. While the use of a continuous function over this space does, in a sense, correspond to a probabilistic interpretation, there exists an exact range of possible wealth levels, which is not necessarily the case when working with a probability density distribution. These approaches can, however, be shown to be equivalent, using a simple transform, as was demonstrated in §4.3.

The simple per capita wealth growth-rate function (4.16) that models the temporal evolution of a wealth distribution over time is related to reality only in mimicking the expected behaviours resulting from three different notions: First, that of increasing relative inequality over time in the absence of redistribution, secondly, that the attained growth rate at a wealth level depends on the total wealth and, finally, that increases in inequality result in decreased growth prospects at fixed wealth levels. A linear function of wealth was chosen for its simplicity and not because this is expected to be related to reality. The deductions made from the analysis in the following chapters therefore indicate potential behaviours resulting from these notions, and not predictions of behaviours which will necessarily follow in reality.

4.7 Chapter summary

Two mathematical models of wealth distribution were derived in this chapter. First, the underlying model assumptions were noted and motivated, after which the mathematical model derivations were carried out. Metrics for characterising the extent of equality present in model solutions were then defined and their equivalence with traditional measures of wealth distribution was demonstrated. This was followed by a brief discussion on the interpretation and use of these metrics. The per capita wealth growth-rate functions to be considered later in this thesis, in the context of the models derived earlier, were then introduced. The existence and uniqueness of the models' solutions were established and it was demonstrated that these solutions, as well as their first derivatives in space, remain nonnegative for all time. The models' limitations and intended use were finally discussed in a formal model apology.

CHAPTER 5

Conditions for bounded inequality

Contents

In this chapter, sufficient conditions are derived for bounded inequality of wealth. The aim is to show analytically, within the context of the models of Chapter 4, that it is always possible to redistribute wealth in a manner that prohibits increases in wealth inequality. In other words, the stabilising nature of the redistribution terms in the models of Chapter 4 can always be made to dominate the instability of the reaction terms (wealth growth functions). First, a number of useful lemmas are established which are used in proofs of subsequent results. Conditions for bounded relative inequality are then considered, and this is followed by the establishment of conditions for bounded absolute inequality. The analytic results of the chapter are finally illustrated numerically.

5.1 Preliminary analysis

The change in mean or total wealth of a distribution is of considerable interest in analyses of economic growth over time. The following Lemma relates the rate of change in the total population wealth to the type of per capita growth-rate function.

Lemma 5.1 *The rate of change in total wealth in* (4.1)–(4.3) *or* (4.4)–(4.5) *is given by*

$$\mu'(t) = \int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda.$$

Proof. By definition,

$$\mu'(t) = \int_{\mathcal{D}} \frac{\partial w}{\partial t}\,\mathrm{d}\lambda,$$

and upon substitution of (4.1),

$$\begin{aligned}
\mu'(t) &= \int_{\mathcal{D}} \left(wf(w) + d\frac{\partial^2 w}{\partial \lambda^2} \right) \mathrm{d}\lambda \\
&= \int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda + d\int_{\mathcal{D}} \frac{\partial^2 w}{\partial \lambda^2}\,\mathrm{d}\lambda \qquad (5.1) \\
&= \int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda + d\left(\left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=0} - \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=1} \right) \qquad (5.2) \\
&= \int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda
\end{aligned}$$

by utilisation of the fundamental theorem of the calculus (see Theorem 3.2) in (5.1) and the boundary conditions (4.2) in (5.2).

Furthermore, in the case of (4.4),

$$\begin{aligned}
\mu'(t) &= \int_{\mathcal{D}} \frac{\partial w}{\partial t}\,\mathrm{d}\lambda \\
&= \int_{\mathcal{D}} wf(w) + d(\mu - w)\,\mathrm{d}\lambda \\
&= \int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda + d\mu - d\int_{\mathcal{D}} w\,\mathrm{d}\lambda \qquad (5.3) \\
&= \int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda
\end{aligned}$$

upon substitution of (4.6) into (5.3). □

The rate of change in relative inequality will be used in arguments related to upward or downward trends in relative inequality later in this chapter.

Lemma 5.2 *The rate of change in relative inequality for the initial-boundary value problem* (4.1)–(4.3) *or the initial value problem* (4.4)–(4.5) *is given by*

$$\overline{V}'(t) = \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda \right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda \right) + \int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - dr(t) \right], \qquad (5.4)$$

where $r(t) = \int_{\mathcal{D}} \left(\frac{\partial w}{\partial \lambda}\right)^2 \mathrm{d}\lambda$ *in the case of* (4.1) *and* $r(t) = V(t)$ *in the case of* (4.4) *represents the effect of redistribution.*

Proof. In the case of both the models (4.1)–(4.3) and (4.4)–(4.5), it follows that

$$\overline{V}'(t) = \frac{-2\mu'}{\mu^3}\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda + \frac{2}{\mu^2}\int_{\mathcal{D}} w\frac{\partial w}{\partial t}\,\mathrm{d}\lambda. \qquad (5.5)$$

So, by using Lemma 5.1 and substituting (4.1) into (5.5), it follows that

$$\begin{aligned}
\overline{V}'(t) &= \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) + \int_{\mathcal{D}} w\left(wf(w) + d\frac{\partial^2 w}{\partial \lambda^2}\right)\mathrm{d}\lambda\right] \\
&= \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) + \int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda + d\int_{\mathcal{D}} w\frac{\partial^2 w}{\partial \lambda^2}\,\mathrm{d}\lambda\right] \qquad (5.6) \\
&= \frac{2}{\mu^2}\Bigg[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) + \int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda \\
&\qquad + d\left(w\frac{\partial w}{\partial \lambda}\bigg|_{\lambda=0}^{\lambda=1} - \int_{\mathcal{D}}\left(\frac{\partial w}{\partial \lambda}\right)^2\mathrm{d}\lambda\right)\Bigg] \qquad (5.7) \\
&= \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) + \int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - d\int_{\mathcal{D}}\left(\frac{\partial w}{\partial \lambda}\right)^2\mathrm{d}\lambda\right]
\end{aligned}$$

in the case of (4.1)–(4.3) upon utilisation of the fundamental theorem of the calculus, as well as integration by parts, in (5.6) and the boundary conditions (4.2) in (5.7).

Furthermore, by using Lemma 5.1 and substituting (4.4) into (5.5), it follows that

$$\begin{aligned}
\overline{V}'(t) &= \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) + \int_{\mathcal{D}} w\left(wf(w) + d(\mu - w)\right)\mathrm{d}\lambda\right] \\
&= \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) + \int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda + d\left(\mu\int_{\mathcal{D}} w\,\mathrm{d}\lambda - \int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right)\right] \qquad (5.8) \\
&= \frac{2}{\mu^2}\Bigg[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) + \int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda \\
&\qquad + d\left(\mu^2 - \int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right)\Bigg] \qquad (5.9) \\
&= \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) + \int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - dV(t)\right]
\end{aligned}$$

upon substitution of (4.6) into (5.8) and (4.11) into (5.9). □

The rates of change in absolute inequality in both the trickle-down and linear redistribution models are to be used in subsequent arguments related to upward or downward trends in absolute inequality. The following result is therefore useful.

Lemma 5.3 *The rate of change in absolute inequality for the initial-boundary value problem* (4.1)–(4.3) *or the initial value problem* (4.4)–(4.5) *is given by*

$$V'(t) = 2\left[\int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda - dr(t)\right],$$

where $r(t) = \int_{\mathcal{D}}\left(\frac{\partial w}{\partial \lambda}\right)^2\mathrm{d}\lambda$ *in the case of* (4.1) *and* $r(t) = V(t)$ *in the case of* (4.4).

Proof. In the case of both the models (4.1)–(4.3) and (4.4)–(4.5), it holds that

$$\begin{aligned}
V'(t) &= 2\int_{\mathcal{D}} (w - \mu)\left(\frac{\partial w}{\partial t} - \mu'\right)\mathrm{d}\lambda \\
&= 2\int_{\mathcal{D}}\left(w\frac{\partial w}{\partial t} - \mu' w - \mu\frac{\partial w}{\partial t} + \mu\mu'\right)\mathrm{d}\lambda. \qquad (5.10)
\end{aligned}$$

By using Lemma 5.1 substituting and (4.1) into (5.10), it follows, performing integration by parts, that

$$\begin{aligned} V'(t) = 2\Bigg[&\int_{\mathcal{D}} w\left(wf(w) + d\frac{\partial^2 w}{\partial\lambda^2}\right)\mathrm{d}\lambda - \left(\int_{\mathcal{D}} w\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right) \\ &- \mu\int_{\mathcal{D}}\left(wf(w) + d\frac{\partial^2 w}{\partial\lambda^2}\right)\mathrm{d}\lambda + \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\Bigg] \qquad (5.11) \\ = 2&\left[\int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda + d\left(\int_{\mathcal{D}} w\frac{\partial^2 w}{\partial\lambda^2}\,\mathrm{d}\lambda + \int_{\mathcal{D}}\frac{\partial^2 w}{\partial\lambda^2}\,\mathrm{d}\lambda\right)\right] \qquad (5.12) \\ = 2&\Bigg[\int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda \\ &+ d\left(w\frac{\partial w}{\partial\lambda}\bigg|_{\lambda=0}^{\lambda=1} - \int_{\mathcal{D}}\left(\frac{\partial w}{\partial\lambda}\right)^2\mathrm{d}\lambda + \frac{\partial w}{\partial\lambda}\bigg|_{\lambda=0} - \frac{\partial w}{\partial\lambda}\bigg|_{\lambda=1}\right)\Bigg]. \qquad (5.13) \end{aligned}$$

Substitution of (4.6) into (5.11) and utilisation of the fundamental theorem of the calculus (see Theorem 3.2) in (5.12) as well as the boundary conditions (4.2) in (5.13) therefore yield

$$V'(t) = 2\left[\int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda - d\int_{\mathcal{D}}\left(\frac{\partial w}{\partial\lambda}\right)^2\mathrm{d}\lambda\right]$$

in the case of (4.1)–(4.3). By using Lemma 5.1 and substituting (4.4) into (5.10) instead, it follows that

$$\begin{aligned} V'(t) = 2\Bigg[&\int_{\mathcal{D}} w[wf(w) + d(\mu - w)]\,\mathrm{d}\lambda - \left(\int_{\mathcal{D}} w\,\mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\right) \\ &- \mu\int_{\mathcal{D}}[wf(w) + d(\mu - w)]\,\mathrm{d}\lambda + \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\Bigg] \qquad (5.14) \\ = 2\Bigg[&\int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda + d\left(\mu\int_{\mathcal{D}} w\,\mathrm{d}\lambda - \int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right) - 2\mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda \\ &- d\left(\mu^2 - \mu\int_{\mathcal{D}} w\,\mathrm{d}\lambda\right) + \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda\Bigg] \qquad (5.15) \\ = 2&\left[\int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda + d\left(\mu^2 - \int_{\mathcal{D}} w^2\,\mathrm{d}\lambda\right)\right] \qquad (5.16) \\ = 2&\left[\int_{\mathcal{D}} w^2 f(w)\,\mathrm{d}\lambda - \mu\int_{\mathcal{D}} wf(w)\,\mathrm{d}\lambda - dV(t)\right] \end{aligned}$$

upon substitution of (4.6) into (5.14) and (5.15), and substitution of (4.11) into (5.16). □

5.2 Conditions for nonincreasing inequality

The per capita wealth growth-rate functions considered in this section are all special cases of (4.16). Substituting (4.16) into (4.1) yields the initial-boundary value problem

$$\begin{aligned} &\text{D.E.} \quad \frac{\partial w}{\partial t} = w(a + bw - c\mu - k\overline{V}) + d\frac{\partial^2 w}{\partial\lambda^2} \quad \lambda \in \mathcal{D},\ t \geq 0, && (5.17) \\ &\text{B.C.} \quad \frac{\partial w}{\partial\lambda}\bigg|_{\lambda=0} = \frac{\partial w}{\partial\lambda}\bigg|_{\lambda=1} = 0,\ t \geq 0, && (5.18) \\ &\text{I.C.} \quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0,\ \lambda \in \mathcal{D}, && (5.19) \end{aligned}$$

whereas substituting (4.16) into (4.4) yields the initial value problem

$$\text{D.E.} \quad \frac{\partial w}{\partial t} = w(a + bw - c\mu - k\overline{V}) + d[\mu(t) - w(\lambda, t)] \quad \lambda \in \mathcal{D},\ t \geq 0, \tag{5.20}$$

$$\text{I.C.} \quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0,\ \lambda \in \mathcal{D}. \tag{5.21}$$

Lemma 5.4 *Let $w(\lambda, t)$ be the unique solution to (4.1)–(4.3) or (4.4)–(4.5). Then the change in the total population wealth over time is given by*

$$\mu'(t) = [a - c\mu - k\overline{V}]\mu + b\int_{\mathcal{D}} w^2 \, \mathrm{d}\lambda.$$

Proof. According to Lemma 5.1 the change in the mean wealth is given by

$$\begin{aligned}\mu'(t) &= \int_{\mathcal{D}} wf(w) \, \mathrm{d}\lambda \\ &= \int_{\mathcal{D}} (aw + bw^2 - c\mu w - k\overline{V}w) \, \mathrm{d}\lambda \\ &= (a - c\mu - k\overline{V})\mu + b\int_{\mathcal{D}} w^2 \, \mathrm{d}\lambda. \end{aligned}$$ □

As a validation of the model logic, it is deduced from Lemma 5.4 that for the simple case of constant per capita growth (where $b = c = k = 0$ and $a \neq 0$) the total population wealth is an exponential function of time. In this simple case, the rate of growth of population wealth is independent of the local distribution of wealth.

A necessary and sufficient condition for decreasing relative inequality over time is now presented.

Proposition 5.1 *(A condition for nonincreasing relative wealth inequality)*
Let $w(\lambda, t)$ be the unique solution to (4.1)–(4.3) or (4.4)–(4.5). Then the relative inequality over $w(\lambda, t)$ is a strictly decreasing function of time at time t if and only if

$$dr > b\left[\int_{\mathcal{D}} w^3 \mathrm{d}\lambda - \frac{1}{\mu}\left(\int_{\mathcal{D}} w^2 \mathrm{d}\lambda\right)^2\right]. \tag{5.22}$$

Proof. If, and only if, (5.22) holds, it follows from Lemmas 5.2 and 5.4 that

$$\begin{aligned}\overline{V}'(t) &= \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} w(a + bw - c\mu - k\overline{V}) \, \mathrm{d}\lambda\right)\left(\int_{\mathcal{D}} w^2 \, \mathrm{d}\lambda\right)\right. \\ &\qquad \left. + \int_{\mathcal{D}} w^2(a + bw - c\mu - k\overline{V}) \, \mathrm{d}\lambda - dr(t)\right] \\ &= \frac{2}{\mu^2}\left[\frac{-1}{\mu}\left(\int_{\mathcal{D}} w^2 \, \mathrm{d}\lambda\right)\left[(a - c\mu - k\overline{V})\mu + b\int_{\mathcal{D}} w^2 \, \mathrm{d}\lambda\right]\right. \\ &\qquad \left. + (a - c\mu - k\overline{V})\int_{\mathcal{D}} w^2 \, \mathrm{d}\lambda + b\int_{\mathcal{D}} w^3 \, \mathrm{d}\lambda - dr(t)\right] \\ &= \frac{2}{\mu^2}\left[b\int_{\mathcal{D}} w^3 \, \mathrm{d}\lambda - \frac{b}{\mu}\left(\int_{\mathcal{D}} w^2 \, \mathrm{d}\lambda\right)^2 - dr(t)\right] \\ &< 0. \end{aligned}$$ □

It is expected that, since relative inequality is invariant under multiplication by a constant, relative inequality should be decreasing (and constant only in the case of total equality) in the presence of redistribution for the case of constant per capita growth. This expectation is verified in the following corollary.

Corollary 5.1 *(A sufficient condition for nonincreasing inequality in a model special case)* *Let $w(\lambda, t)$ be the unique solution to* (4.1)–(4.3) *or* (4.4)–(4.5)*. If $b = c = k = 0$, $a \neq 0$ and $d > 0$, then $\overline{V}'(t) \leq 0$.*

Proof. It follows from Proposition 5.1 that for $b = 0$, relative inequality is nonincreasing over time, that is $\overline{V}'(t) < 0$, if $dr(t) > 0$. In the case of (5.17), $r(t) = \int_{\mathcal{D}} \left(\frac{\partial w}{\partial \lambda}\right)^2 \mathrm{d}\lambda$, which is nonnegative by definition, and strictly positive if $V(t) > 0$, since positive variance implies that the spatial derivative of w is nonzero somewhere in $\mathcal{D}$. Furthermore, in the case of (5.20), $r(t) = V(t)$ and the desired result follows immediately. □

An easily verifiable sufficient condition for decreasing relative inequality of wealth in the initial-boundary value problem (5.17)–(5.19) may be derived from Proposition 5.1.

Corollary 5.2 *(A sufficient condition for nonincreasing wealth inequality)* *Let $w(\lambda, t)$ be the unique solution to* (4.1)–(4.3)*. If*

$$d[w_{\max}(t) - w_{\min}(t)]^2 > b\left[w_{\max}^3(t) - \frac{w_{\min}^4(t)}{w_{\max}(t)}\right] \tag{5.23}$$

at some time t, then $\overline{V}'(t) < 0$.

Proof. It follows from the Cauchy-Buniakovskii-Schwarz inequality (see Corollary 3.1) and by Proposition 4.3 that

$$r(t) = \int_{\mathcal{D}} \left(\frac{\partial w}{\partial \lambda}\right)^2 \mathrm{d}\lambda \geq \left[\int_{\mathcal{D}} \left(\frac{\partial w}{\partial \lambda}\right) \mathrm{d}\lambda\right]^2 = [w(1,t) - w(0,t)]^2 = [w_{\max}(t) - w_{\min}(t)]^2.$$

Suppose that (5.23) holds. Then, by Propositions 4.3 and 5.1,

$$\begin{aligned}
\overline{V}'(t) &= \frac{2}{\mu^2}\left[b\int_{\mathcal{D}} w^3 \,\mathrm{d}\lambda - \frac{b}{\mu}\left(\int_{\mathcal{D}} w^2 \,\mathrm{d}\lambda\right)^2 - dr(t)\right] \\
&\leq \frac{2}{\mu^2}\left[bw^3(1,t) - b\frac{w^4(0,t)}{w(1,t)} - d[w_{\max}(t) - w_{\min}(t)]^2\right] \\
&= \frac{2}{\mu^2}\left[bw_{\max}^3(t) - b\frac{w_{\min}^4(t)}{w_{\max}(t)} - d[w_{\max}(t) - w_{\min}(t)]^2\right] \\
&< 0.
\end{aligned}$$

□

The significant implication of Corollary 5.2 is that in the context of the models in this thesis, a redistribution rate which ensures nonincreasing relative inequality can be determined using only the extremal values of the wealth distribution.

Absolute inequality is now considered. Although the primary focus in this thesis is on relative inequality, it is expected that the treatment of absolute inequality may contribute to the reader's understanding of the difference between the two concepts (hence clarifying the precise meaning of relative inequality), and therefore be beneficial.

Proposition 5.2 *(A condition for nonincreasing absolute wealth inequality)*
Let $w(\lambda, t)$ be the unique solution to (4.1)–(4.3) *or* (4.4)–(4.5). *Then the absolute inequality over $w(\lambda, t)$ is a strictly decreasing function of time at some time t if and only if*

$$dr > (a - c\mu - k\overline{V})V + b\left(\int_{\mathcal{D}} w^3 \,\mathrm{d}\lambda - \mu \int_{\mathcal{D}} w^2 \,\mathrm{d}\lambda\right). \tag{5.24}$$

Proof. If, and only if, (5.24) is satisfied, it follows form Lemmas 5.3 and 5.4 that

$$\begin{aligned}
V'(t) &= 2\left[\int_{\mathcal{D}} w^2(a + bw - c\mu - k\overline{V}) \,\mathrm{d}\lambda - \mu \int_{\mathcal{D}} w(a + bw - c\mu - k\overline{V}) \,\mathrm{d}\lambda - dr(t)\right] && (5.25)\\
&= 2\left[(a - c\mu - k\overline{V}) \int_{\mathcal{D}} w^2 \,\mathrm{d}\lambda + b\int_{\mathcal{D}} w^3 \,\mathrm{d}\lambda - (a - c\mu - k\overline{V})\mu^2 - b\mu \int_{\mathcal{D}} w^2 \,\mathrm{d}\lambda - dr(t)\right] \\
&= 2\left[(a - c\mu - k\overline{V})\left(\int_{\mathcal{D}} w^2 \,\mathrm{d}\lambda - \mu^2\right) + b\left(\int_{\mathcal{D}} w^3 \,\mathrm{d}\lambda - \mu \int_{\mathcal{D}} w^2 \,\mathrm{d}\lambda\right) - dr(t)\right] && (5.26)\\
&= 2\left[(a - c\mu - k\overline{V})V + b\left(\int_{\mathcal{D}} w^3 \,\mathrm{d}\lambda - \mu \int_{\mathcal{D}} w^2 \,\mathrm{d}\lambda\right) - dr(t)\right] \\
&< 0 && (5.27)
\end{aligned}$$

upon substitution of (4.11) into (5.26). □

It should be noted that the parameters a, c and k explicitly influence the change in absolute inequality over time, although this is not the case with relative inequality. Furthermore, finding a nontrivial sufficient condition for decreasing absolute inequality without utilising a distributional metric, such as the variance, cannot be done as easily as in the case of relative inequality.

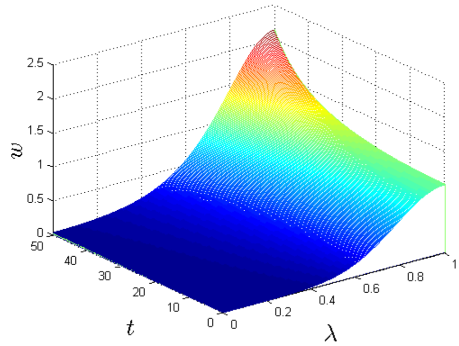
5.3 Numerical examples

The conditions for decreasing relative inequality in §5.2 are illustrated in this section by means of numerical examples.

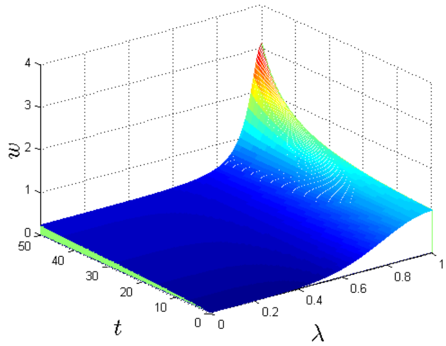
First, the result of Proposition 5.1 is illustrated by plotting the normalised variance of wealth over time as well as both sides of the inequality (5.22), for instances of both the trickle-down model and the linear redistribution model. This is done for the simple per capita wealth growth-rate function (4.14), shown in Figure 5.1. It may be verified that the location in time of the extremal points in Figures 5.22 (c) correspond to the crossing of the plots in Figures 5.22 (b). As expected, relative inequality decreases over time in the solutions to both the trickle-down model and the linear redistribution model while $dr(t)$ is greater than the right-hand side of (5.22).

It is demonstrated next that the sufficient condition for decreasing relative inequality of Corollary 5.2 is not necessary. The same linear per capita wealth growth-rate function is employed as in the previous example and a large redistribution rate is chosen to ensure that (5.23) is satisfied initially. Relative inequality is decreasing over the time interval $[0, 5]$ in Figure 5.2(c), yet (5.23) is satisfied only on the first half of this interval in Figure 5.2(b). It may therefore be concluded that (5.23) need not be satisfied for relative inequality to decrease over time.

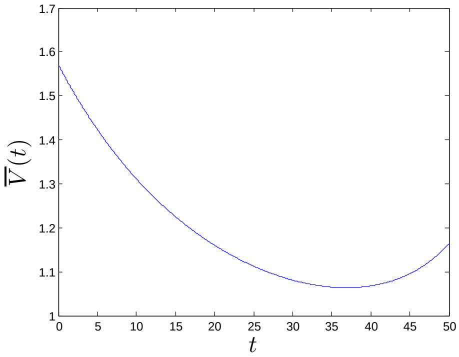
Finally, the necessary and sufficient condition for decreasing absolute inequality of Proposition 5.2 is illustrated for both models. The solution variance, as well the left and right-hand sides of the inequality (5.24), are plotted over time in Figure 5.3. It is apparent that equality of the left and right-hand side components of (5.24) in Figure 5.3(b) correspond to an extremal



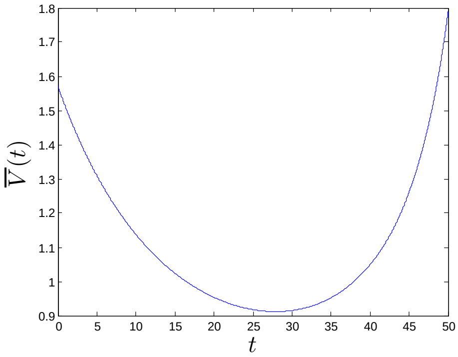
(a) *Trickle-down model*



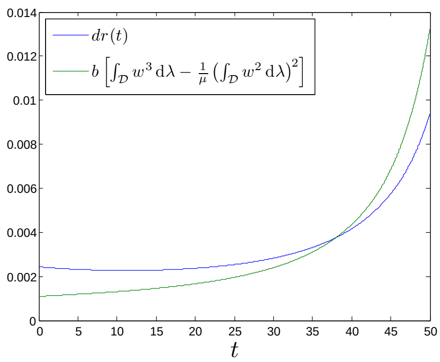
(a) *Linear redistribution model*



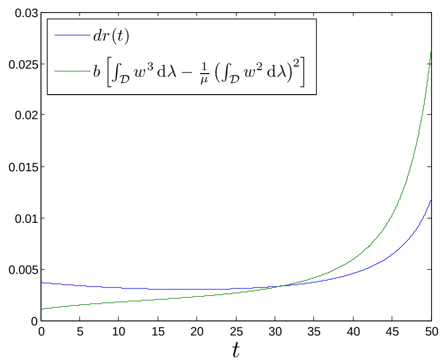
(b) *Trickle-down model*



(b) *Linear redistribution model*

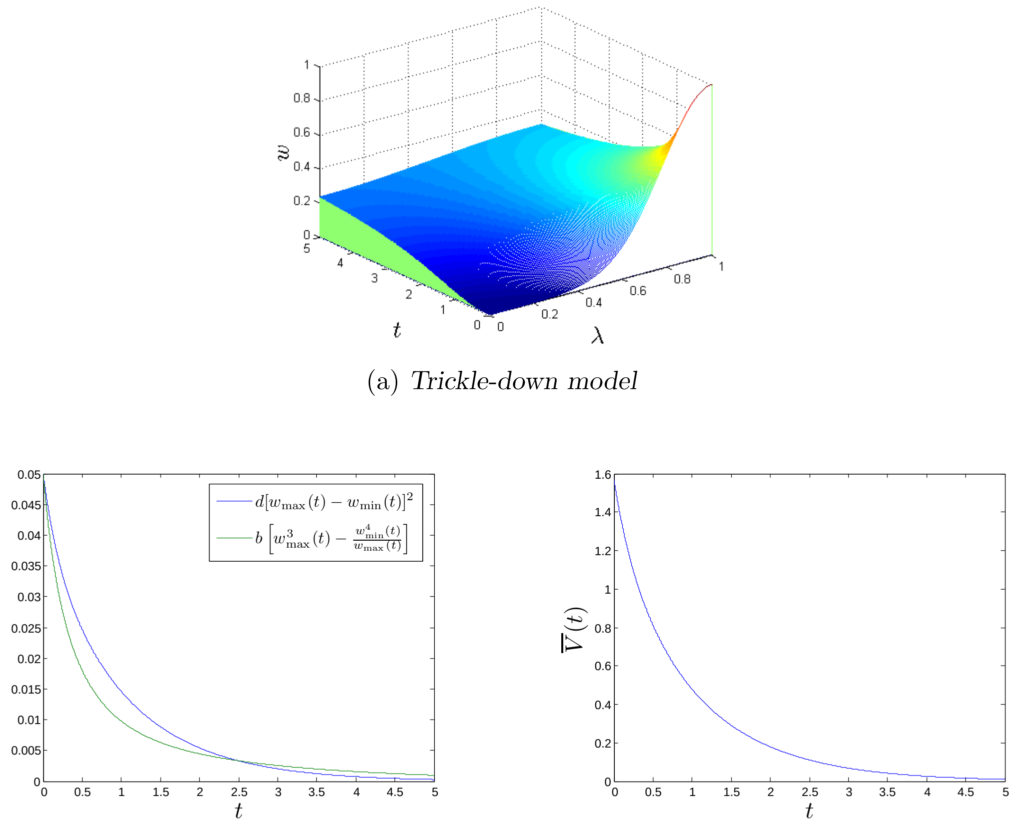


(c) *Trickle-down model*



(c) *Linear redistribution model*

FIGURE 5.1: *An example of necessary and sufficient conditions for decreasing relative inequality in the models (4.1)–(4.3) and (4.4)–(4.5). The figures are based on numerical approximations of the model solutions obtained by applying the method (3.29) in the case of the trickle-down model and Euler's method (3.27) in the case of the linear redistribution model, in both cases with step sizes* $\triangle t = 0.01$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$*, and assuming the initial condition* $w_0(\lambda) = \exp(-10(\lambda-1)^2)$ *and parameter values* $a = -0.025$ *and* $b = 0.05$*. In the case of the trickle-down model,* $d = 1.2 \times 10^{-3}$*, while in the case of the linear redistribution model,* $d = 0.03$*.*



(a) *Trickle-down model*

(b) *Inequality* (5.23) *components*

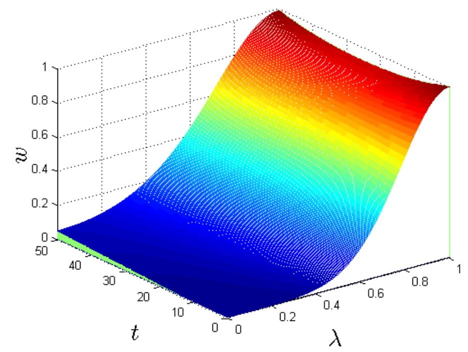
(c) *Relative inequality*

FIGURE 5.2: *An example of the sufficient condition for decreasing relative inequality the model* (4.1)–(4.3). *The figures are based on numerical approximations of the solution to* (4.1)–(4.3) *obtained by applying the method* (3.29) *with step sizes* $\triangle t = 1 \times 10^{-3}$ *and* $\triangle\lambda = 0.0125$, *and assuming the initial condition* $w_0(\lambda) = \exp(-10(\lambda - 1)^2)$ *and parameter values* $a = -0.025$, $b = 0.05$ *and* $d = 0.05$.

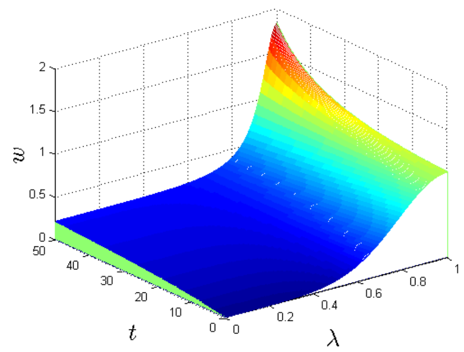
point in the absolute inequality in Figure 5.3(c). Furthermore, while $dr(t)$ dominates the right-hand side of (5.24), absolute inequality decreases over time, as predicted by Proposition 5.2.

5.4 Chapter summary

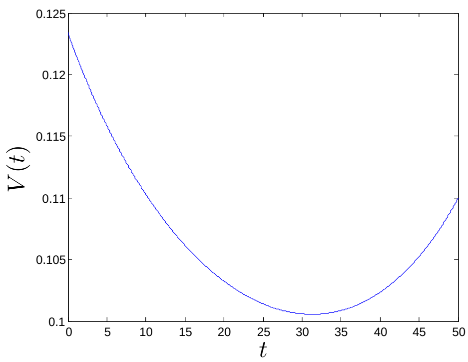
The models derived in Chapter 4 were validated in this chapter by analysing their solution behaviour for a constant per capita wealth growth rate. An expression for the change in total wealth over time was derived for all per capita wealth growth-rate function types considered. A necessary and sufficient condition for decreasing relative inequality was established. An easily verifiable sufficient condition for decreasing relative inequality was derived from this necessary and sufficient condition which depends only the extremal values of the wealth distribution. A necessary and sufficient condition for decreasing absolute inequality was also established and the differences between this condition and that for relative inequality was noted. Numerical examples illustrating some of the analytic results established in this chapter were finally presented.



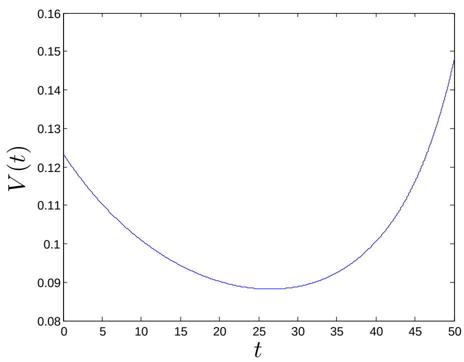
(a) *Trickle-down model*



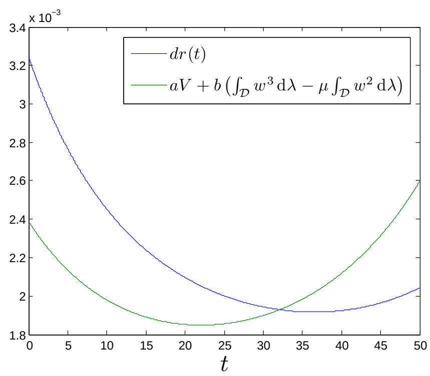
(a) *Linear redistribution model*



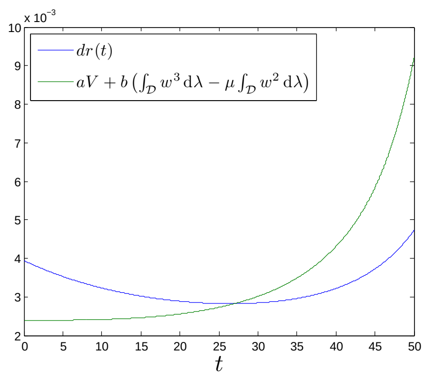
(b) *Trickle-down model*



(b) *Linear redistribution model*



(c) *Trickle-down model*



(c) *Linear redistribution model*

FIGURE 5.3: *An example of necessary and sufficient conditions for decreasing absolute inequality in the models* (4.1)–(4.3) *and* (4.4)–(4.5). *The figures are based on numerical approximations of the model solutions obtained by applying the method* (3.29) *in the case of the trickle-down model and Euler's method* (3.27) *in the case of the linear redistribution model, in both cases with step sizes* $\triangle t = 0.01$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$, *and assuming the initial condition* $w_0(\lambda) = \exp(-10(\lambda-1)^2)$ *and parameter values* $a = -0.025$ *and* $b = 0.05$. *In the case of the trickle-down model,* $d = 1.6 \times 10^{-3}$, *while in the case of the linear redistribution model,* $d = 0.032$.

CHAPTER 6

Long-time solution behaviour

Contents

The long-time solution behaviour of the models of Chapter 4 is investigated in this chapter. The per capita wealth growth-rate functions of §4.4 are considered in order of increasing complexity. For the simple case of a linear per capita wealth growth rate, the relationship between solution persistence and the redistribution rate is investigated. The stability of constant equilibrium solutions is analysed and the existence of a spatially nonconstant equilibrium solution to the trickle-down model is established. For the mean-dependent per capita wealth growth-rate function, the stability of equilibrium solutions and the sizes of the basins of attraction of stable equilibrium solutions are investigated numerically. The relationship between the characteristics of the nonconstant equilibrium solution and the redistribution rate is considered. Conjectures based on the findings of the numerical experiments are put forward. For the mean and inequality-dependent per capita wealth growth-rate function, possible solution behaviours are demonstrated numerically. Throughout, observations related to possible implications for economic systems are noted, to be elaborated on in the subsequent chapter.

6.1 Linear per capita wealth growth

In this section, the special case of (4.16) where $c = k = 0$ is considered; this corresponds to the growth function (4.14). In the case of the trickle-down model (4.1)–(4.3), substituting (4.14) into (4.1) gives rise to the initial-boundary value problem

$$\text{D.E.} \quad \frac{\partial w}{\partial t} = w(a + bw) + d\frac{\partial^2 w}{\partial \lambda^2} \quad \lambda \in \mathcal{D},\ t \geq 0, \tag{6.1}$$

$$\text{B.C.} \quad \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=0} = \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=1} = 0,\ t \geq 0, \tag{6.2}$$

$$\text{I.C.} \quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0,\ \lambda \in \mathcal{D}. \tag{6.3}$$

In the case of the global redistribution model (4.4)–(4.5), substitution of (4.14) into (4.4) yields the initial value problem

$$\text{D.E.} \quad \frac{\partial w}{\partial t} = w(a + bw) + d[\mu(t) - w(\lambda, t)] \quad \lambda \in \mathcal{D},\ t \geq 0 \tag{6.4}$$

$$\text{I.C.} \quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0,\ \lambda \in \mathcal{D}. \tag{6.5}$$

Since the focus here is on linear per capita growth rates with $b > 0$, and since wealth is nonnegative by Proposition 4.2, the population per capita growth rate for a positive total wealth is not less than a. If a is nonnegative, then more redistribution therefore results in a lower positive population wealth growth rate. If $a < 0$, and a portion of the population possesses wealth levels below $w_c = -a/b$, however, then the population wealth growth rate may be negative, causing the population wealth to tend to zero over time. It is not unreasonable to assume that there exists such a critical wealth level, below which an entity's wealth is unable to grow. Persons whose living expenses equal or exceed their total income, or businesses whose operating expenses exceed their income, are examples of entities below such a critical level. Furthermore, if wealth growth in the models of Chapter 4 is viewed as growth in value (therefore growth measured in monetary value discounted according to inflation), then it is reasonable to expect that certain entities should experience a decline in wealth over time, since not all persons or businesses experience monetary growth greater than inflation at all times.

It is of interest to establish exactly under which conditions the persistence of population wealth may be threatened. Cases where $a < 0$ and $w_0(0) < -a/b < w_0(1)$ must therefore be considered. If all entities have less wealth than the critical wealth (that is, $-a/b > w_0(1)$), then the entire population will experience negative growth for all time, while if all entities have more wealth than the critical wealth (that is, $-a/b < w_0(0)$), then the entire population will experience positive growth for all time. It is shown first that, for any initial distribution satisfying $w_0(0) < -a/b < w_0(1)$, a redistribution rate d exists below which increasing total population wealth is guaranteed as $t \to \infty$. In other words, if a proportion of entities, no matter how small, possess wealth above the critical wealth, then their wealth positions can increase over time, given that the net influence of redistribution is smaller than this growth.

Proposition 6.1 *If $a < 0$ and $b > 0$ in (6.1)–(6.3) or (6.4)–(6.5), and $w_0(1) > -a/b$, then there exists a redistribution rate $d_\infty > 0$ such that $\mu(t) \to \infty$ as $t \to \infty$ for all $d \in [0, d_\infty)$.*

Proof. The evolution of wealth over time in both (6.1)–(6.3) and (6.4)–(6.5) is governed by the equation

$$\frac{\partial w}{\partial t} = aw + bw^2 + dR(w).$$

It follows from the existence of classical solutions to (6.1)–(6.3) and (6.4)–(6.5) that

$$|R(w)| \leq M$$

on some time interval $[0, T_{\max})$, where $T_{\max}$ and M are positive real numbers.

The inequality

$$\frac{\partial w}{\partial t} > aw(\lambda, t) + bw^2(\lambda, t) - dM \tag{6.6}$$

then represents a lower bound on $\partial w/\partial t$ at a given time $t \in [0, T_{\max})$ and position λ. If

$$d < \frac{1}{M}(aw_c + bw_c^2), \tag{6.7}$$

where $w_c = -a/b$ denotes the critical wealth, then it follows that $\partial w/\partial t > 0$ for all $w > w_c$. According to Euler's method,

$$w(\lambda, t + \triangle t) = w(\lambda, t) + \triangle t \frac{\partial w(\lambda, t)}{\partial t} + \frac{\triangle t^2}{2} \frac{\partial^2 w(\lambda, t + \alpha \triangle t)}{\partial t^2},$$

for some $\alpha \in [0, 1]$, and so there exists a range of values for $\triangle t$, such that $w(\lambda, t+\triangle t) > w(\lambda, t)$ for all $\lambda \in (w^{-1}(w_c, t), 1]$, where w^{-1} is the inverse function of w such that $w(\lambda, t) = w^{-1}(w(\lambda, t), t)$. Hence, the positive lower bound on $\partial w/\partial t$ on the interval $\lambda \in (w_0^{-1}(w_c), 1]$ increases over any time interval when (6.7) is satisfied. It therefore follows that $\int_{w_0^{-1}(w_c)}^{1} w(\lambda, t)\, \mathrm{d}\lambda \to \infty$ as $t \to \infty$, and since $w(\lambda, t) > 0$, it follows that $\mu(t) \to \infty$ as $t \to \infty$. □

In the cases under consideration, it is therefore always possible to select a redistribution rate such that population wealth persists over time. It is now established exactly when the population wealth can tend to zero. First, however, a property of the redistributive term is established which is to be used in the proof of the subsequent proposition.

Lemma 6.1 *In solutions to* (6.1)–(6.3) *or* (6.4)–(6.5) *the redistribution term satisfies* $R(w) \leq 0$ *at* $\lambda = 1$.

Proof. If $V(t) = 0$, which implies that w is constant over $\mathcal{D}$ at time t, then $R(w) = 0$ at time t. In the case of (6.1), the desired result follows since the first and second derivatives of a constant are zero. In the case of (6.4), the mean of a constant function is equal to the function, and therefore the difference between the function and its mean is zero. If $V(t) > 0$, then according to Proposition 4.3 the maximum wealth in $\mathcal{D}$ at any time t is given by $w(1, t)$. It follows from the continuity of w that $\partial^2 w/\partial \lambda^2 \leq 0$ at such a maximum, in the case of (6.1). Since the maximum value of a function is not less than its mean, a nonpositive value is obtained when subtracting the maximum from the mean and therefore the desired result also follows in the case of (6.4). □

Proposition 6.2 *If* $a < 0$ *and* $b > 0$ *in* (6.1)–(6.3) *or* (6.4)–(6.5), *and* $\int_{\mathcal{D}} w_0\, \mathrm{d}\lambda < -a/b$, *then there exists a redistribution rate* d_0 *such that* $\mu(t) \to 0$ *as* $t \to \infty$ *for all* $d > d_0$.

Proof. Suppose, at an arbitrary time $t = t_a > 0$, that the maximum wealth satisfies $w(1, t_a) < w_c$. Then the change in population wealth, given by

$$\mu'(t) = \int_{\mathcal{D}} w f(w)\, \mathrm{d}\lambda = \int_{\mathcal{D}} w(a + bw)\, \mathrm{d}\lambda,$$

is negative at $t = t_a$ since $a + bw < 0$ for all w. Furthermore, the change in maximum wealth is also negative, since $R(w)$ is nonpositive at the maximum wealth according to Lemma 6.1. Since $f'(w) > 0$, a decrease in wealth causes a decrease in the per capita wealth growth rate f, and it therefore follows that the total wealth continues to decrease for all $t > t_a$.

Suppose at time $t = 0$ that $\int_{\mathcal{D}} w_0\, \mathrm{d}\lambda < w_c$ and that $w_0(1) > w_c$. Then, if

$$\left.\frac{\partial w}{\partial t}\right|_{\lambda=1} < \frac{w_c - w_0(1)}{t_a} \tag{6.8}$$

for all $t \in [0, t_a)$, it follows that $w(1, t_a) < w_c$. Let the maximum value of $R(w)$ at $\lambda = 1$ over all $t \in [0, t_a)$ be m ($m < 0$ according to Lemma 6.1). By selecting

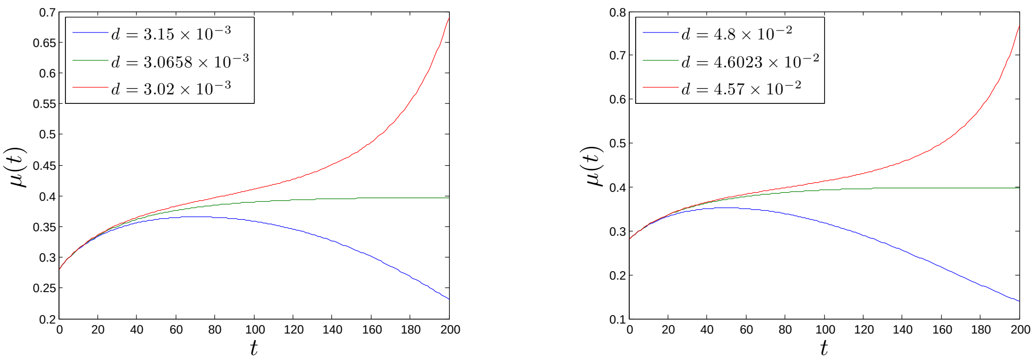
$$d > \frac{-1}{m} \left(a w_0(1) + b w_0^2(1) + \frac{w_0(1) - w_c}{t_a} \right),$$

the inequality (6.8) is satisfied for all $t \in [0, t_a)$. □

It has been established that there exist scenarios in the context of linear per capita wealth growth rates where the persistence of population wealth hinges on the redistribution rate. If the mean wealth is less than the critical wealth required for positive growth, $\mu(t) < w_c$, and therefore the growth function evaluated at the mean wealth is negative (that is, $f(\mu(t)) < 0$), it follows that excessive redistribution will result in a continuous decline in population wealth. If a portion of entities possess more than the critical wealth, such that $w(1,t) > 0$, it has also been established that the population wealth may increase continually when the extent of redistribution is sufficiently limited. Hence it follows that there exists a critical redistribution rate $d_c \in (d_\infty, d_0)$, which causes the mean wealth to tend to a constant value as $t \to \infty$. The following corollary therefore follows immediately from Propositions 6.1 and 6.2.

Corollary 6.1 *Suppose that $a < 0$, $\mu(0) < w_c$ and $w_0(1) > w_c$, where $w_c = -a/b$. Then there exists a $d_c \in (d_\infty, d_0)$ such that $\mu(t) \to \mu_e$ as $t \to \infty$, and therefore $\partial w/\partial t \to 0$ as $t \to \infty$, where $\mu_e \in (0, w_c]$ is a constant.*

The critical redistribution rate d_c of Corollary 6.1 can easily be approximated numerically using an iterative search technique. Examples of such critical redistribution rates are illustrated in Figure 6.1 for the parameter values and initial condition in Table 6.1.



(a) *Trickle-down model* (b) *Linear redistribution model*

FIGURE 6.1: *Population wealth over time for linearly increasing per capita wealth growth rates and wealth distributed around a critical wealth level. The figures are based on numerical approximations of the model solutions obtained by applying the method (3.29) in the case of the trickle-down model and Euler's method (3.27) in the case of the linear redistribution model, in both cases with step sizes $\triangle t = 0.01$ and $\triangle\lambda = 1.25 \times 10^{-3}$, and assuming the initial conditions and parameter values in Table 6.1.*

Parameter	Value
a	-2×10^{-2}
b	5×10^{-2}
$-a/b$	0.4
$w_0(\lambda)$	$\exp(-10(\lambda - 1)^2)$

TABLE 6.1: *Parameter values used to illustrate possible solution behaviours for linear per capita wealth growth rates when $a < 0$ and $w_0(0) < -a/b < w_0(1)$.*

If the persistence of population wealth is a requirement (as it should be), the existence of a critical redistribution rate imposes a limit on the maximum permissible redistribution rate. When w_0 is increased (uniformly or proportionally), and f is unchanged, then the value of d_c in Corollary 6.1 increases, as evident in Table 6.2. The associated trends in mean wealth are shown in Figure 6.2. This phenomenon is intuitive, since in the context of this growth function, an increase in wealth causes an increase in the growth of wealth. An economy with better growth prospects can therefore redistribute to a greater extent, without inducing an adverse downward trend over time in the total population wealth.

Initial condition	Trickle-down d_c	Linear redistribution d_c
w_0	3.06514×10^{-3}	4.60199×10^{-2}
$w_0 + 0.05$	4.81074×10^{-3}	6.38550×10^{-2}
$w_0 + 0.1$	1.43330×10^{-2}	1.72102×10^{-1}
$1.1w_0$	4.47938×10^{-3}	6.14207×10^{-2}
$1.2w_0$	7.05309×10^{-2}	9.06162×10^{-2}

TABLE 6.2: *Estimated critical redistribution rates for the parameters in Table 6.1.*

The following observation related to the *Robin Hood paradox* follows from the existence of the critical redistribution rate of Corollary 6.1.

Observation 6.1 *Consider several hypothetical societies, each having a different total wealth level and each pursuing two economic aims. The first aim is nonnegative growth of total wealth and therefore the persistence of population wealth over time, and the second aim is the pursuit of increased economic equality within each society. These societies may then be described as egalitarian in a pragmatic sense. If the existence of increasing per capita wealth growth rates and a critical wealth requirement for positive growth are accepted, then the extent of redistribution present in economies with worse growth prospects (with mean wealth levels below the critical wealth) will be less than that present in economies with more promising growth prospects.*

If it is furthermore accepted that the growth prospects of very economically unequal societies are worse than those of more equal societies, as suggested in the literature reviewed in §2.2, then the Robin Hood paradox can follow necessarily under the assumptions outlined in Observation 6.1.

The positive equilibrium solutions in the context of linear increasing per capita wealth growth rates, achieved when $d = d_c$, exhibit unstable behaviour. This is to be expected since a positive (or negative) perturbation results in an increase (or a decrease) in $\partial w/\partial t$. While reference to this steady state solution is useful as the boundary of separation between two distinct solution behaviours, it is not of practical interest. Unstable steady state solutions do not manifest themselves in reality, or even in numerical approximations. The stability (or otherwise) of the constant equilibrium solutions to (6.1)–(6.3) is now established analytically.

Let w_e denote an equilibrium solution to (6.1)–(6.3) and let $v(\lambda, t)$ be a function representing a perturbation from this equilibrium. Any solution may then be expressed in the form

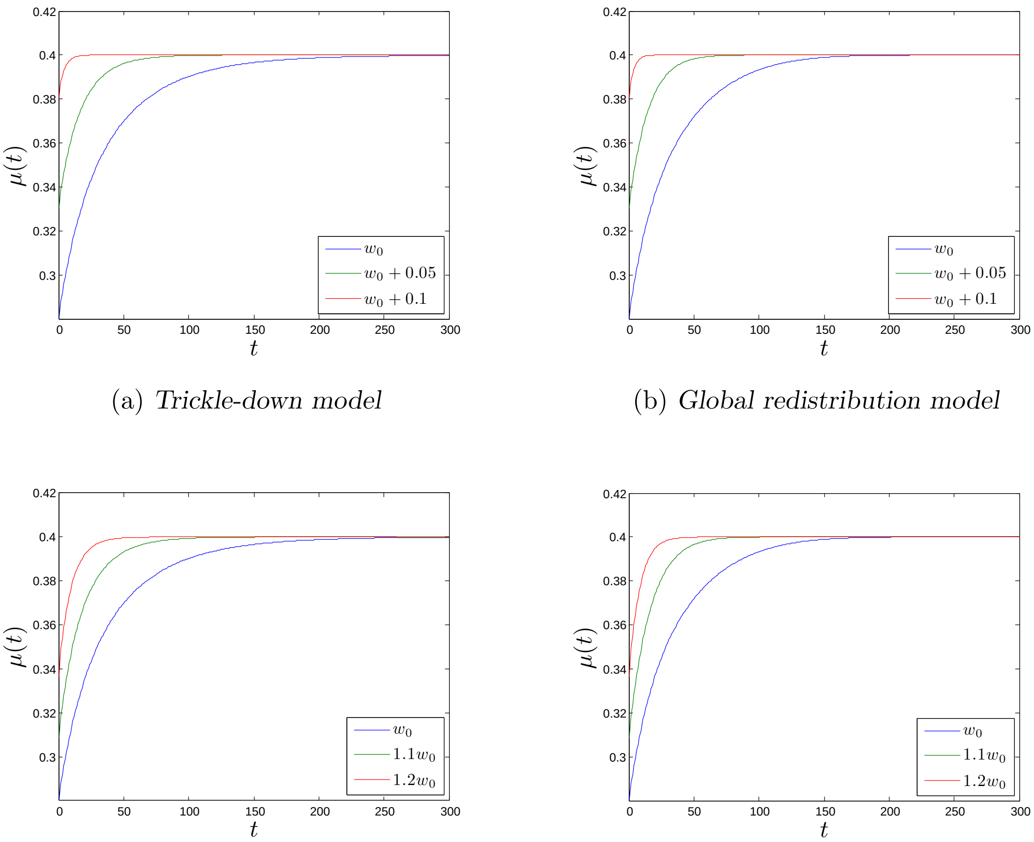
$$w(\lambda, t) = w_e(\lambda) + \epsilon v(\lambda, t), \tag{6.9}$$

where ϵ is a positive real number. Upon substitution of (6.9) into (6.1), it follows that

$$\frac{\partial}{\partial t}(w_e + \epsilon v) = a(w_e + \epsilon v) + b(w_e + \epsilon v)^2 + d\frac{\partial^2}{\partial \lambda^2}(w_e + \epsilon v)$$

and so

$$\epsilon\frac{\partial v}{\partial t} = \epsilon\left(av + 2bw_e v + d\frac{\partial^2 v}{\partial \lambda^2}\right) + b\epsilon^2 v^2,$$



(a) *Trickle-down model* (b) *Global redistribution model*

(c) *Trickle-down model* (d) *Global redistribution model*

FIGURE 6.2: *Population wealth over time for linearly increasing per capita wealth growth rates, wealth distributed around a critical wealth level and a critical redistribution rate. The figures are based on numerical approximations of the model solutions obtained by applying the method (3.29) in the case of the trickle-down model and Euler's method (3.27) in the case of the linear redistribution model, in both cases with step sizes $\triangle t = 0.01$ and $\triangle\lambda = 1.25 \times 10^{-3}$, and assuming the initial conditions and parameter values in Table 6.1.*

which may be linearised to approximate solutions to (6.1)–(6.3) at equilibrium points. If this is done, the perturbation v satisfies

$$\frac{\partial v}{\partial t} = av + 2bw_e v + d\frac{\partial^2 v}{\partial \lambda^2}. \tag{6.10}$$

Imposing the boundary conditions (6.3) on (6.9), that is

$$\frac{\partial}{\partial \lambda}(w_e + \epsilon v)\bigg|_{\lambda=0} = \frac{\partial}{\partial \lambda}(w_e + \epsilon v)\bigg|_{\lambda=1} = 0,$$

yields

$$\frac{\partial v}{\partial \lambda}\bigg|_{\lambda=0} = \frac{\partial v}{\partial \lambda}\bigg|_{\lambda=1} = 0. \tag{6.11}$$

Separation of variables may be applied to (6.10). Let $v(\lambda, t) = X(\lambda)T(t)$. Then (6.10) becomes

$$XT' = (a + 2bw_e)XT + dX''T,$$

which may be rearranged as

$$\frac{T'}{T} = a + 2bw_e + d\frac{X''}{X} = c_1, \tag{6.12}$$

where c_1 is a constant.

Therefore, $T(t) = c_2 e^{c_1 t}$, where c_2 is a constant, and

$$X'' + \underbrace{\frac{1}{d}(a + 2bw_e - c_1)}_{\varphi^2} X = 0.$$

For w_e a constant, the cases $\varphi^2 < 0$, $\varphi = 0$ and $\varphi > 0$ are considered separately. If $\varphi^2 < 0$, then $X(\lambda) = Ae^{\varphi t} + Be^{-\varphi t}$, where A and B are constants. Imposing the boundary conditions (6.11) yields the trivial solution since $A = B = 0$.

If $\varphi = 0$ (*i.e.* $c_1 = a + 2bw_e$), then $X(\lambda) = A\lambda + B$. The boundary conditions require that $A = 0$, and hence that $X(\lambda) = B$. The general solution is then given by

$$v(\lambda, t) = Ce^{(a+2bw_e)t}, \tag{6.13}$$

where $C = c_2 B$.

Finally, if $\varphi > 0$, then $X(\lambda) = A\cos(\varphi\lambda) + B\sin(\varphi\lambda)$. Imposing the boundary conditions yields $B = 0$ and $\varphi = n\pi$ for $n \in \{1, 2, \dots\}$. Therefore, $c_1 = a + 2bw_e - dn^2\pi^2$. The general solution is then given by

$$v(\lambda, t) = \sum_{n=1}^{\infty} A_n e^{(a+2bw_e - dn^2\pi^2)t} \cos(n\pi\lambda). \tag{6.14}$$

The linear stability of a constant equilibrium solution is established by interpreting the sign of c_1 at the equilibrium solution. It is apparent from (6.13) and (6.14) that for $a < 0$, the equilibrium solution $w_e(\lambda) = 0$ is linearly stable since $c_1 < 0$. The solution $w_e(\lambda) = -a/b$ is, however, unstable, since (6.13) simplifies to $v(\lambda, t) = Ce^{-at}$, which grows arbitrarily large over time when $a < 0$.

The existence of a third equilibrium solution to (6.1)–(6.3), which is not constant over space, is now established.

Proposition 6.3 *For any $a < 0$, $b > 0$ and $w(0, t) \in (0, -a/b)$, there exists a nonconstant steady state solution to the model* (6.1)–(6.3).

Proof. A steady state solution to (6.1)–(6.3) is also a solution to the two-point boundary value problem

$$\text{D.E.} \quad \frac{\mathrm{d}^2 w}{\mathrm{d}\lambda^2} = \frac{-w}{d}(a + bw) \tag{6.15}$$

$$\text{B.C.} \quad \left.\frac{\mathrm{d}w}{\mathrm{d}\lambda}\right|_{\lambda=0} = \left.\frac{\mathrm{d}w}{\mathrm{d}\lambda}\right|_{\lambda=1} = 0, \tag{6.16}$$

and a solution to this boundary value problem is a steady state solution to (6.1)–(6.3) if it is nondecreasing, according to Proposition 4.3.

By inspection, $w = 0$ and $w = -a/b$ are solutions to (6.15)–(6.16), and steady state solutions to (6.1)–(6.3). The existence of a third solution is now established. Let $\mathrm{d}w/\mathrm{d}\lambda = q$. Then

$$\frac{\mathrm{d}q}{\mathrm{d}w} q = \frac{-w}{d}(a + bw)$$

and so

$$\int q \, \mathrm{d}q = \frac{-1}{d} \int aw + bw^2 \, \mathrm{d}w,$$

from which it follows that

$$q^2 = \frac{-1}{d} \left(aw^2 + \frac{2bw^3}{3} \right) + c_3,$$

where c_3 is a constant. According to the boundary condition $\mathrm{d}w/\mathrm{d}\lambda = q = 0$ at $w(0) = w_{\text{min}}$,

$$c_3 = \frac{-1}{d} \left(aw_{\text{min}}^2 + \frac{2bw_{\text{min}}^3}{3} \right).$$

Therefore,

$$q^2 = \frac{-1}{d} \left(aw^2 + \frac{2bw^3}{3} - aw_{\text{min}}^2 - \frac{2bw_{\text{min}}^3}{3} \right),$$

and so it follows that

$$q = \sqrt{\frac{-1}{d} \left(aw^2 + \frac{2bw^3}{3} - aw_{\text{min}}^2 - \frac{2bw_{\text{min}}^3}{3} \right)},$$

because of the nonnegativity of $\mathrm{d}w/\mathrm{d}\lambda$. Let $h(w) = q^2$. Then

$$h'(w) = \frac{-2w}{d}(a + bw).$$

Hence the cubic function h has extremal points at $w = 0$ and $w = -a/b$. Choosing w_{min}, which is a root of $h(w) = 0$ by definition, in the interval $(0, -a/b)$ therefore implies that w_{min} is the central root of three distinct real roots of $h(w)$, as illustrated in Figure 6.3. This also implies that $h(w)$ has two positive roots and one negative root. Upon factorisation,

$$h(w) = \frac{-1}{d}(w - w_{\text{min}}) \left(\frac{2bw^2}{3} + \left(a + \frac{2bw_{\text{min}}}{3} \right) w + \left(a + \frac{2bw_{\text{min}}}{3} \right) w_{\text{min}} \right).$$

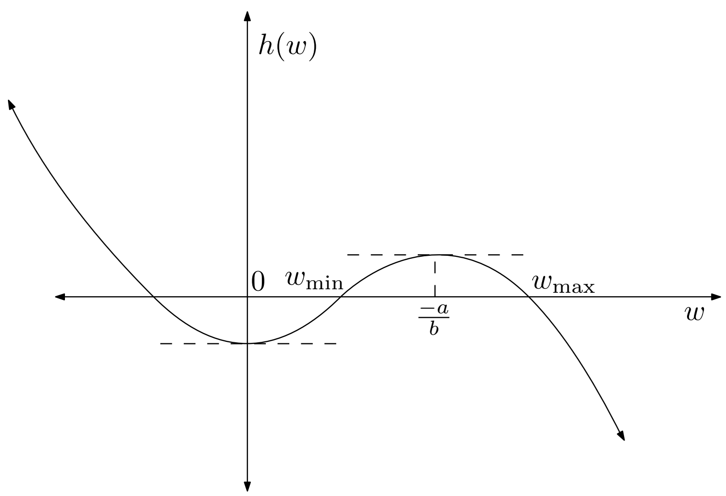


FIGURE 6.3: *The roots of $h(w)$.*

Furthermore, q is real-valued on the interval $[w_{\text{min}}, w_{\text{max}}]$, where w_{max} is the largest root of $h(w) = 0$. Utilising the quadratic formula,

$$w_{\text{max}} = \frac{-3b}{4}\left(a + \frac{2bw_{\text{min}}}{3} - \sqrt{a^2 - \frac{4bw_{\text{min}}}{3}(a + bw_{\text{min}})}\right).$$

Let $\mathcal{L}$ be the curve q on $[w_{\text{min}}, w_{\text{max}}]$ and define

$$\begin{aligned}
\lambda^* &= \int_{\mathcal{L}} \mathrm{d}\lambda \\
&= \int_{\mathcal{L}} \left(\frac{\mathrm{d}w}{\mathrm{d}\lambda}\right)^{-1} \frac{\mathrm{d}w}{\mathrm{d}\lambda}\,\mathrm{d}\lambda \\
&= \int_{w_{\text{min}}}^{w_{\text{max}}} \frac{1}{q}\,\mathrm{d}w.
\end{aligned}$$

If $\lambda^* = 1$, then the boundary conditions are satisfied since λ^* represents the only point at which $\mathrm{d}w/\mathrm{d}\lambda$ becomes zero for $w > w_{\text{min}}$. This can always be achieved by selecting a suitable value for d. Setting

$$\int_{w_{\text{min}}}^{w_{\text{max}}} \frac{1}{q}\,\mathrm{d}w = 1,$$

it follows that

$$\int_{w_{\text{min}}}^{w_{\text{max}}} \left[\frac{-1}{d}\left(aw^2 + \frac{2bw^3}{3} - aw_{\text{min}}^2 - \frac{2bw_{\text{min}}^3}{3}\right)\right]^{-\frac{1}{2}} \mathrm{d}w = 1$$

and solving for d yields

$$d = \left[\int_{w_{\text{min}}}^{w_{\text{max}}} \left(aw^2 + \frac{2bw^3}{3} - aw_{\text{min}}^2 - \frac{2bw_{\text{min}}^3}{3}\right)^{-\frac{1}{2}} \mathrm{d}w\right]^{-2}. \tag{6.17}$$

The elliptic integral[1] in (6.17) is convergent. Therefore, for any $a < 0$, $b > 0$, $w_{\text{min}} \in (0, -a/b)$ and d defined as in (6.17), there exists a nondecreasing, non-constant solution to the boundary value problem (6.15)–(6.16) and therefore a steady state solution to (6.1)–(6.3). □

Attention is now turned to the long-time solution behaviour of the model (6.4)–(6.5). First, the behaviour of the differential equation is treated separately from the initial conditions.

Proposition 6.4 *Equation* (6.4) *has, in addition to the trivial zero solution, one positive constant equilibrium solution and infinitely many two-valued step-function equilibrium solutions for* $a < 0$, $b > 0$ *and* $d > 0$.

Proof. By inspection and from the definition of a steady state for (6.4),

$$aw + bw^2 + d(\mu - w) = 0, \tag{6.18}$$

it follows that $w = 0$ and $w = -a/b$ are solutions to (6.4). Since (6.18) is a quadratic equation in w and $\mu(t) = \mu_e$ is a constant at a steady state, the equation can have at most of two

[1] An elliptic integral is an integral of the form $\int R(S, \lambda)\,\mathrm{d}\lambda$, where R is a rational function containing at least one odd power of S and S^2 is a cubic or quadratic function of λ with no repeated roots [1].

distinct roots. It is now shown that step-function equilibrium solutions exist to (6.4) for every $\mu_e \in (0, -a/b)$. Rearranging (6.18) as

$$bw^2 + (a-d)w + d\mu_e = 0, \tag{6.19}$$

it follows that two distinct real solutions to (6.18) exist if $(a-d)^2 - 4bd\mu_e > 0$, that is, if

$$\mu_e < \frac{(a-d)^2}{4bd}. \tag{6.20}$$

A steady state to (6.4) can only exist for $\mu_e \leq -a/b$. It is now demonstrated that this, in turn, also ensures that (6.20) is satisfied. By definition $0 \leq (d+a)^2$, and upon expanding the square and subtracting $4ad$,

$$-4ad \leq d^2 - 2ad + a^2,$$

and since both d and b are positive, this can be written as

$$\frac{-a}{b} \leq \frac{(d-a)^2}{4bd}.$$

For any $\mu_e \in (0, -a/b)$, the steady state function values are therefore

$$w_1 = \frac{1}{2b}\left(d - a - \sqrt{(a-d)^2 - 4bd\mu_e}\right)$$

and

$$w_2 = \frac{1}{2b}\left(d - a + \sqrt{(a-d)^2 - 4bd\mu_e}\right).$$

Assuming an increasing step function, a steady state solution to (6.4) is therefore given by

$$w(\lambda) = \begin{cases} w_1 & \text{if } 0 \leq \lambda \leq \frac{\mu_e - w_2}{w_1 - w_2} \\ w_2 & \text{if } \frac{\mu_e - w_2}{w_1 - w_2} < \lambda \leq 1 \end{cases} \tag{6.21}$$

for any $\mu_e \in (0, -a/b)$. □

The initial conditions (6.5) ensure that the model (6.4)–(6.5) possesses continuous solutions, and therefore the step function steady state solutions (6.21) are not solutions of (6.4)–(6.5).

Corollary 6.2 *The model* (6.4)–(6.5)*, with* $a < 0$, $b > 0$ *and* $d > 0$ *possesses only two steady state solutions, given by* $w = 0$ *and* $w = -a/b$. *Of these, the former is stable, while the latter is unstable.*

In summary, the solutions of the models (6.1)–(6.3) and (6.4)–(6.5) either tend to zero or grow without bound over time, since all positive equilibrium solutions are unstable.

6.2 Mean-dependent per capita wealth growth

In this section, possible solution behaviours of the wealth distribution and redistribution models are investigated for a specific mean-dependent per capita wealth growth-rate function. Substitution of the per capita growth-rate function (4.15) into (4.1)–(4.3) gives rise to the initial-boundary value problem

$$\text{D.E.} \quad \frac{\partial w}{\partial t} = w(a + bw - c\mu) + d\frac{\partial^2 w}{\partial \lambda^2} \quad \lambda \in \mathcal{D},\ t \geq 0, \tag{6.22}$$

$$\text{B.C.} \quad \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=0} = \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=1} = 0,\ t \geq 0, \tag{6.23}$$

$$\text{I.C.} \quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0,\ \lambda \in \mathcal{D}, \tag{6.24}$$

while substitution of (4.15) into (4.4)–(4.5) yields the initial value problem

$$\text{D.E.} \quad \frac{\partial w}{\partial t} = w(a + bw - c\mu) + d[\mu(t) - w(\lambda, t)] \quad \lambda \in \mathcal{D},\ t \geq 0 \tag{6.25}$$

$$\text{I.C.} \quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0,\ \lambda \in \mathcal{D}. \tag{6.26}$$

It can be demonstrated numerically that the simple addition of the $c\mu$-term to the linear per capita wealth growth-rate function creates stable steady state solutions for a large class of parameter combinations. If $a > 0$, then solutions to (6.22)–(6.24) or (6.25)–(6.26) cannot tend to zero. This is evident from the reaction term $w(a + bw - c\mu)$, since as w becomes small, a dominates, and so the reaction term is positive for small values of w when $a > 0$.

If $b \geq c$, then solutions to (6.22)–(6.24) or (6.25)–(6.26) grow unbounded. Since per capita growth rates are increasing, the slowest possible growth rate for the total population wealth occurs at total equality of wealth. This corresponds to the solution of the ODE initial value problem

$$\frac{\mathrm{d}w}{\mathrm{d}t} = aw + (b - c)w^2, \tag{6.27}$$

$$w(0) = \mu(0) = w_0, \tag{6.28}$$

since $\mu = w$ when w is constant. It is therefore easily seen that even the slowest possible growth corresponds to solutions growing unbounded if $c \leq b$. Letting $b < c$ creates a stable equilibrium solution to (6.27) at $w = a/(c - b)$. Hence there exist constant stable equilibrium solutions to (6.22)–(6.24) and (6.25)–(6.26), since the redistribution term ensures regression to the mean provided that perturbations are sufficiently small. Since the redistribution term in (6.25)–(6.26) allows only constant equilibrium solutions, which are neither mathematically interesting nor practically interpretable, this model is not analysed further.

The stability of the nonconstant equilibrium solution to (6.22)–(6.24) is of interest. The existence of such a solution is guaranteed by Proposition 6.3. If an equilibrium solution to (6.1)–(6.3) has a mean μ_e, then the same function is an equilibrium solution to (6.22)–(6.24) provided that $a - c\mu_e$ in (6.22) replaces the value of a in (6.1). Analytically expressing this nonconstant equilibrium solution or analysing its stability properties is, however, difficult. The following conjecture is therefore substantiated numerically.

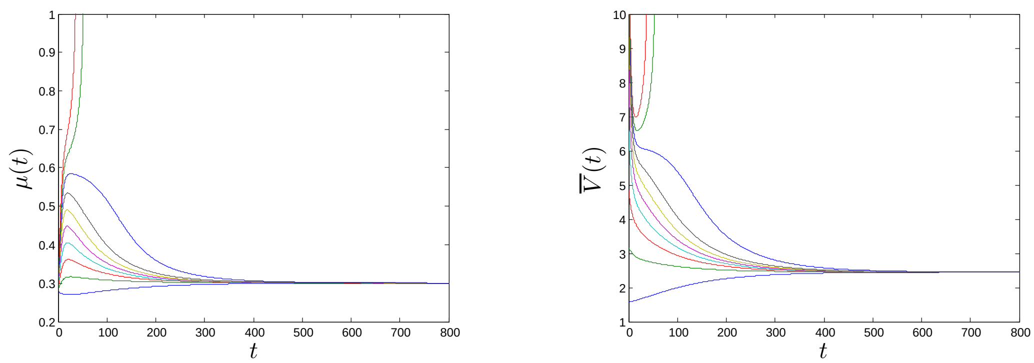
Conjecture 6.1 *If $a > 0$, $c > b > 0$ and $d > d_\infty > 0$, where d_∞ is some lower bound required for stability at the nonconstant equilibrium solution to (6.22)–(6.24), then there exists a nonempty set $\mathcal{W}$ of $\mathcal{C}^2$ functions defined on $\mathcal{D}$ such that, for all $w_0 \in \mathcal{W}$, the solution w to (6.22)–(6.24) tends to the nonconstant equilibrium solution as $t \to \infty$.*

The left-hand side portion of a Gaussian distribution is once again used to construct initial conditions of the form (6.24). The mean and variance of such a distribution uniquely determine an initial condition. Given a set of parameters and initial conditions which converge to a stable equilibrium (such as for the parameters in Table 6.3 together with $w_0(\lambda) = \exp(-(\lambda-1)^2)$), two types of deviation in the initial conditions are investigated separately. First, the mean is kept constant while the normalised variance is increased. In Figure 6.4 it may be seen that, for a certain fixed mean wealth, the extent of inequality of an initial condition can determine whether the associated solution converges to the equilibrium solution. Since solutions cannot tend to zero in this case, either convergence to an asymptotically stable positive equilibrium or unbounded growth is expected. More unequal distributions cause a larger rate of growth of population wealth. Initial conditions which are more equally distributed than a given initial condition which converges to the equilibrium are therefore expected also to lead to solutions which converge to the same equilibrium solution. This expectation is made precise in the following conjecture.

Conjecture 6.2 *Let $\mathcal{X}$ be the set of all possible initial conditions* (6.24) *defined by* $w_0(\lambda) = k_1 \exp(-k_2(x-1)^2)$ *for any* $k_1, k_2 > 0$. *If a given initial condition* w_0^* *of this form, with mean* μ^* *and normalised variance* $\overline{V}^*$, *is in* $\mathcal{W} \cap \mathcal{X}$ *(where* $\mathcal{W}$ *is as defined in Conjecture 6.1), then all* $w \in \mathcal{X}$ *with* $\mu = \mu^*$ *and* $\overline{V} \leq \overline{V}^*$ *are in* $\mathcal{W}$.

Parameter	Value
a	0.1
b	5×10^{-2}
c	0.5
d	5×10^{-4}

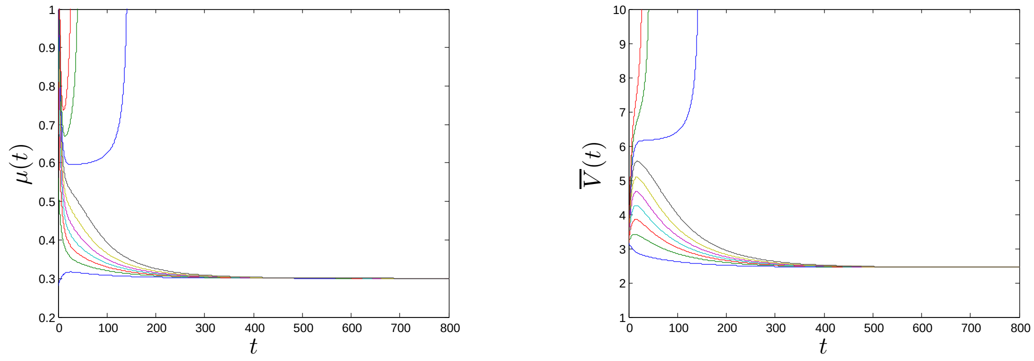
TABLE 6.3: *Parameter values used to illustrate basin of attraction of equilibrium solutions to* (6.22)–(6.24).



(a) *Mean wealth over time*   (b) *Normalised variance over time*

FIGURE 6.4: *Temporal evolution of wealth for initial left-half Gaussian distributions with differing levels of inequality and the same mean for the parameter values in Table 6.3. The figures are based on numerical approximations of solutions to the model* (6.22)–(6.24) *obtained by applying the method* (3.29) *with step sizes* $\triangle t = 0.01$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$.

Next the normalised variance is fixed and the mean wealth is adjusted. From Figure 6.5 it is apparent that, for a fixed normalised variance, the mean wealth of an initial condition can determine whether the associated solution converges to the equilibrium solution. It is therefore concluded that whether or not a function is a member of $\mathcal{W}$ depends on both its shape and vertical scale.



(a) *Mean wealth over time* (b) *Normalised variance over time*

FIGURE 6.5: *Temporal evolution of wealth for initial left-half Gaussian distributions with differing means and the same normalised variance for the parameter values in Table 6.3. The figures are based on numerical approximations to the model* (6.22)–(6.24) *solutions by applying the method* (3.29) *with step size* $\triangle t = 0.01$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$.

Conjecture 6.3 *If $w_0^* \in \mathcal{X}$, $w_0^* \in \mathcal{W}$ and w_0^* has a mean μ^* and normalised variance $\overline{V}^*$, then all $w \in \mathcal{X}$ for which $\overline{V} = \overline{V}^*$ and $\mu \leq \mu^*$ are in $\mathcal{W}$.*

If Conjectures 6.1–6.3 hold, then any function that is dominated in terms of its mean and normalised variance by a function in $\mathcal{X} \cap \mathcal{W}$ (in the Pareto sense for a maximisation problem[2]), is also in $\mathcal{X} \cap \mathcal{W}$.

The *size* of $\mathcal{X} \cap \mathcal{W}$ can therefore in some sense be expressed in terms of the pairs of means and normalised variances of functions in $\mathcal{X}$ which are on the boundary of the basin of attraction of the equilibrium solution. It is expected that this size should increase with an increase in the rate of redistribution. The greater the value of d, the greater the perturbation required from an initial condition inside the basin of attraction so as to allow unstable growth in the reaction term to dominate redistributive effects.

Conjecture 6.4 *Let w_0^* be an initial condition of the form* (6.24) *and suppose $\{a, b, c, d\}$ is a set of positive parameters associated with* (6.22)*, such that, for all $w \in \mathcal{X}$ with corresponding mean $\mu > \mu^*$ and normalised variance $\overline{V} \geq \overline{V}^*$, or $\mu \geq \mu^*$ and $\overline{V} > \overline{V}^*$, w does not converge to the equilibrium solution (w_0^* may then thought of as being 'on the boundary' of $\mathcal{W} \cap \mathcal{X}$). Instead, there exist initial conditions $w' \in \mathcal{X}$ with $\mu' > \mu^*$ and $\overline{V}' \geq \overline{V}^*$, or $\mu' \geq \mu^*$ and $\overline{V}' > \overline{V}^*$, for which the corresponding solutions to* (6.22)–(6.24) *converge to an equilibrium solution corresponding to the parameter set $\{a, b, c, d'\}$ where $d' > d$.*

Figure 6.6 depicts the mean wealth and normalised variance of initial conditions in $\mathcal{X}$ which converge to an equilibrium solution related to the parameters in Table 6.3, for three different values of d. It is apparent that an increase in d results in an associated increase in the size of the region of means and normalised variances related to elements in $\mathcal{X} \cap \mathcal{W}$.

Thus far the investigation in this section has been centred on establishing the stability of nonconstant equilibrium solutions to (6.22)–(6.24), and the effect of varying rates of trickle-down redistribution on the relative size of the basin of attraction to a nonconstant equilibrium solution. Attention is now turned away from the initial conditions and to the effect that such

[2]Meaning that the two measures of the dominated function are both less than or equal to those of the function in $\mathcal{W}$, and one is strictly smaller.

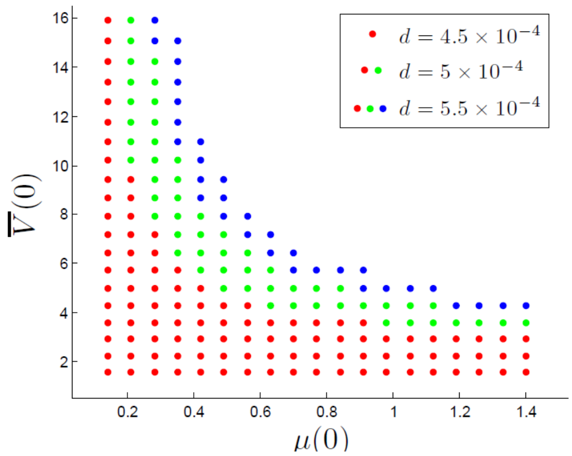


FIGURE 6.6: *The means and normalised variances of selected elements of $\mathcal{X} \cap \mathcal{W}$ for the parameter set specified in Table 6.3, but for varying values of d. The figure is based on numerical approximations of solutions to the model* (6.22)–(6.24) *obtained by applying the method* (3.29) *with step sizes $\triangle t = 0.01$ and $\triangle\lambda = 1.25 \times 10^{-3}$ up to time $t = 2000$, and observing whether unbounded growth above the constant equilibrium occurs.*

parameter variation has on the characteristics of the nonconstant equilibrium solution itself. The parameter values in Table 6.3 are again assumed and estimates of the mean wealth and normalised variance at the equilibrium are recorded for different redistribution rates. The results are depicted in Figure 6.7. The mean wealth at the nonconstant equilibrium solution tends toward the constant equilibrium solution $w = a/(c-b)$ (which equals $0.\dot{2}$ in this case), as shown in Figure 6.7(b). As may be expected, the normalised variance tends to zero as d increases, as shown in Figure 6.7(d).

Conjecture 6.5 *Let $a, b, c, d > 0$ be a parameter set for which there exists a nonconstant equilibrium solution to* (6.22)–(6.24). *Then the mean wealth at the nonconstant equilibrium satisfies $\mu_e > a/(c-b)$. Furthermore, the mean wealth of the nonconstant equilibrium solution to* (6.22)–(6.24) *associated with the parameter set $a, b, c, d' > 0$ and $d' > d$ is $\mu'_e < \mu_e$.*

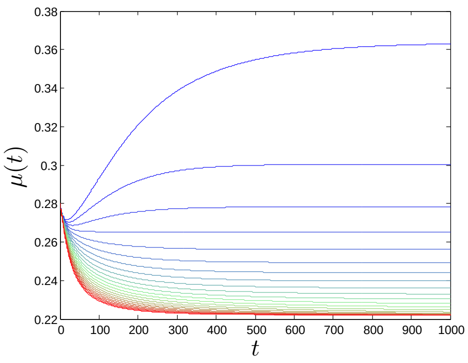
It then follows that the mean wealth at the nonconstant equilibrium solution can be made arbitrarily close to the constant solution $w = a/(c-b)$.

Conjecture 6.6 *Given positive values of a, b and c for which an equilibrium solution to* (6.22)–(6.24) *exists, there exists, for any $\epsilon > 0$, a value of d for which $a/(c-b) < \mu_e < a/(c-b) + \epsilon$.*

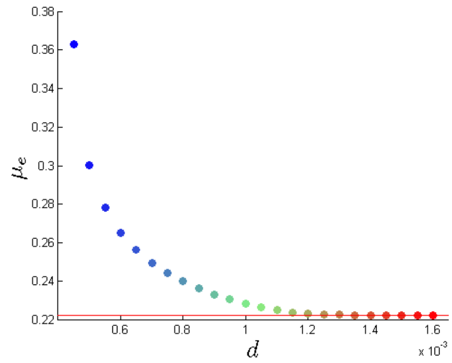
The implications of the conjectures in this section are captured in the following remark.

Observation 6.2 *In the presence of increasing per capita wealth growth rates, competition resulting in decreased growth prospects at a fixed wealth level when the population wealth increases and sufficient trickle-down redistribution, a stable, nonconstant equilibrium wealth distribution exists. The stability of this equilibrium increases with greater trickle-down redistribution effects, but the mean equilibrium wealth decreases with greater trickle-down redistribution.*

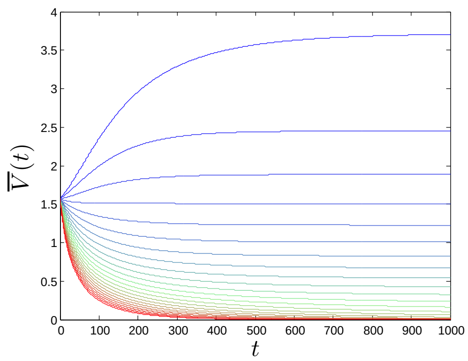
There are two significant, and perhaps counter intuitive, insights to be gained from the conjectures and observations of this section. The first is that a stable nonconstant equilibrium wealth distribution exists, although the reaction term in the model is unstable in the sense that it is an increasing function of wealth everywhere. Allowing an increasing linear per capita wealth



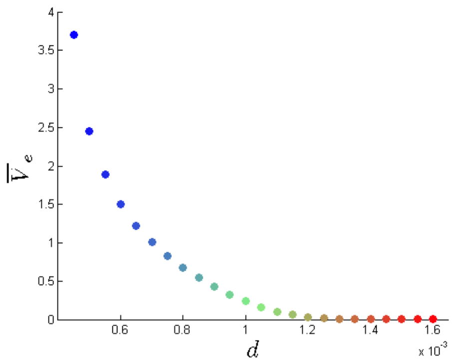
(a) *Equilibrium mean wealth for different values of d*



(b) *Mean wealth over time corresponding to the different values of d in (a)*



(c) *Equlibrium normalised variance for different values of d*



(d) *Normalised variance over time corresponding to the different values of d in (c)*

FIGURE 6.7: *The relationship between the redistribution rate, the mean wealth and the normalised variance at the nonconstant equilibrium solution to* (6.22)–(6.24) *for the parameter values in Table 6.3. The figures are based on numerical approximations of solutions to the model* (6.22)–(6.24) *obtained by applying the method* (3.29) *with step sizes* $\triangle t = 5 \times 10^{-3}$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$.

growth-rate function to translate along the wealth axis, with translations proportional to the mean wealth, is sufficient to render stable nonconstant equilibrium solutions. The implication of this insight is that a system of this type need not be extremely fine-tuned (as no social system can be) in order to be stable. There exists a large range of parameters and corresponding sets of initial conditions that lead to stable equilibrium solutions.

The second insight is the fact that the mean wealth at a nonconstant equilibrium solution is greater than the constant equilibrium solution. The implication is that even in a world of identical agents, the average agent is better off when wealth is not equally distributed. Greater inequality is associated with a larger mean equilibrium wealth. The largest possible mean equilibrium wealth is therefore associated with the highest possible level of inequality and the lowest possible rate of redistribution, for an equilibrium solution corresponding to a given parameter set. This latter state then forms the boundary of the basin of attraction, and is therefore not stable.

6.3 Mean and inequality-dependent per capita wealth growth

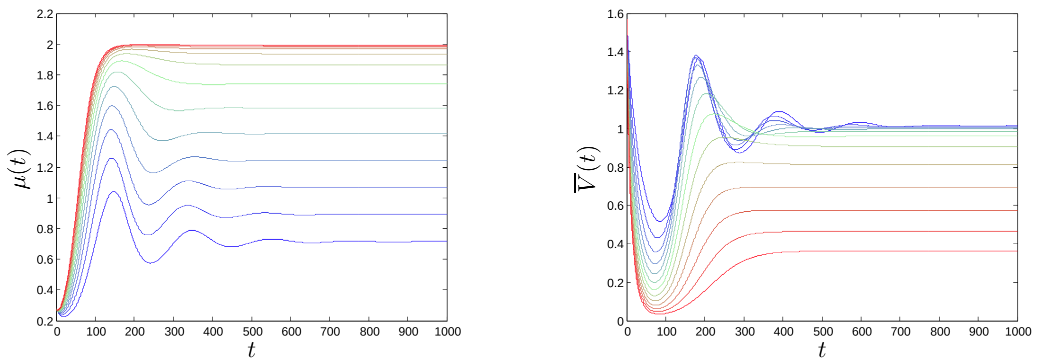
In this section, possible solution behaviours of the model (4.1)–(4.3) with the per capita growth-rate function (4.16) are demonstrated by means of numerical simulations. The initial-boundary value problem

$$\text{D.E.}\quad \frac{\partial w}{\partial t} = w(a + bw - c\mu - k\overline{V}) + d\frac{\partial^2 w}{\partial \lambda^2} \quad \lambda \in \mathcal{D},\ t \geq 0, \tag{6.29}$$

$$\text{B.C.}\quad \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=0} = \left.\frac{\partial w}{\partial \lambda}\right|_{\lambda=1} = 0,\ t \geq 0, \tag{6.30}$$

$$\text{I.C.}\quad w(\lambda, 0) = w_0(\lambda) \geq 0,\ w_0'(\lambda) \geq 0,\ \lambda \in \mathcal{D} \tag{6.31}$$

is considered. Two distinct types of solution behaviour are demonstrated using the parameter set labelled 1 in Table 6.4, for various redistribution rates and two values of k. It is apparent from Figure 6.8 that stable positive equilibrium solutions exist, since the total wealth and normalised variance both tend to constant values over time. For low redistribution rates, vanishing oscillations are observed in both the mean wealth and normalised variance.



(a) *Mean wealth over time for different values of d*

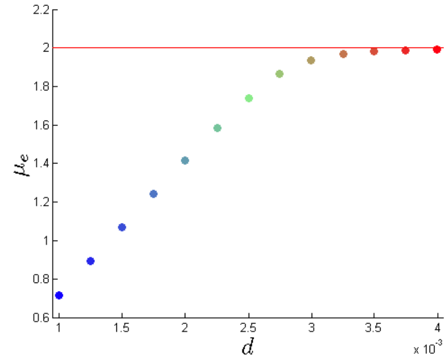
(b) *Normalised variance over time for different values of d*

FIGURE 6.8: *The relationship between the redistribution rate, the mean wealth and the normalised variance at nonconstant equilibrium solutions to* (6.29)–(6.31) *for* $k = 4.9 \times 10^{-2}$, *parameter set 1 and* $d \in \{1, 1.25, 1.5, \ldots, 4\} \times 10^{-3}$. *The figures are based on numerical approximations of solutions to the model* (6.29)–(6.31) *obtained by applying the method* (3.29) *with step sizes* $\triangle t = 5 \times 10^{-3}$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$, *and assuming the initial condition* $w_0(\lambda) = \exp(-10(\lambda - 1)^2)$.

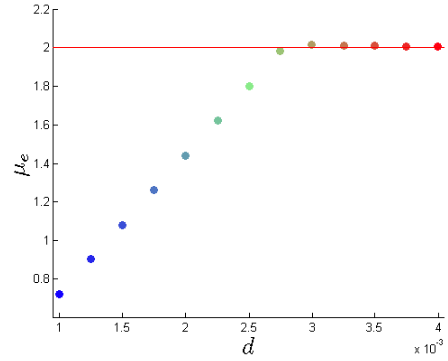
Parameter	Set 1	Set 2	Set 3
a	5×10^{-2}	0.4129×10^{-2}	0.1×10^{-2}
b	2.5×10^{-2}	5.38×10^{-2}	5×10^{-2}
c	5×10^{-2}	0.5519	6.7×10^{-2}
k	—	3.91×10^{-2}	0.1
d	—	—	3×10^{-4}

TABLE 6.4: *Parameter sets used to illustrate solution behaviours for mean and inequality-dependent per capita wealth growth rates.*

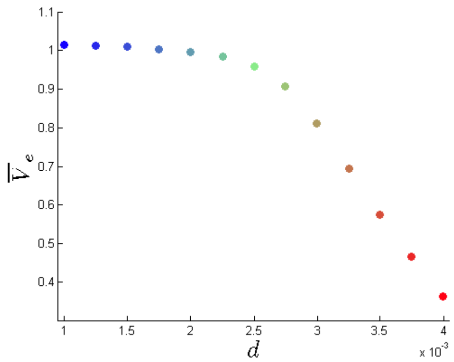
In figure 6.9(a), the relationship between the mean wealth at the equilibrium solution and the rate of redistribution is reversed from that of the previous section (see Figure 6.7(a)). The mean



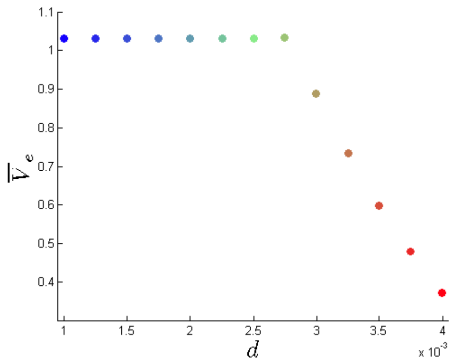
(a) *Equilibrium mean wealth for different values of* d *and* $k = 4.9 \times 10^{-2}$



(b) *Equilibrium mean wealth for different values of* d *and* $k = 4.85 \times 10^{-2}$



(c) *Equlibrium normalised variance for different values of* d *and* $k = 4.9 \times 10^{-2}$

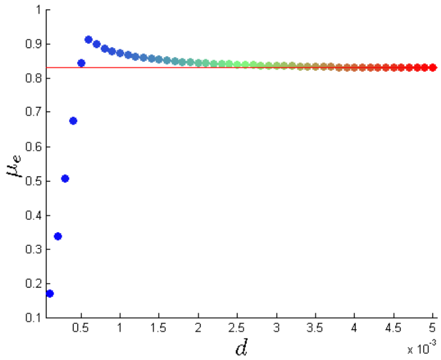


(d) *Equlibrium normalised variance for different values of* d *and* $k = 4.85 \times 10^{-2}$

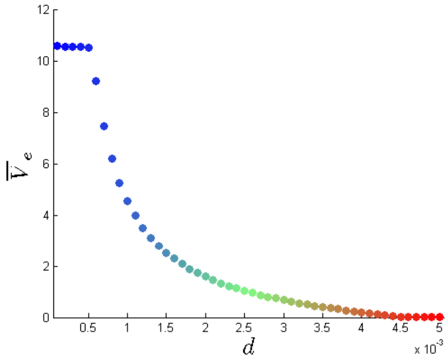
FIGURE 6.9: *The relationship between the redistribution rate, the mean wealth and the normalised variance at nonconstant equilibrium solutions resulting from parameter set 1 in Table 6.4. The figures are based on numerical approximations of solutions to the model* (6.22)–(6.24) *obtained by applying the method* (3.29) *with step sizes* $\triangle t = 5 \times 10^{-3}$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$.

equilibrium wealth is always less than the constant equilibrium (still given by $w = a/(c-b)$), and increases, tending toward the constant solution, as d increases. This rather unsatisfactory result suggests that more redistribution is always preferred, yielding total equality as the best possible state, and providing no real new insights. Fascinating solution behaviour however occurs when k is decreased slightly. In Figure 6.9(b), the mean equilibrium wealth rises above the constant equilibrium at around $d \approx 3 \times 10^{-3}$, and then tends toward $w = a/(c-b)$ from above as d increases. This suggests that there exists a unique, finite redistribution rate d at which a maximum mean equilibrium wealth is achieved. It is also remarkable that, as d increases, this maximum mean wealth occurs after a rather abrupt decrease in the normalised variance. The largest possible mean is therefore not achieved at the highest possible level of inequality associated with an equilibrium solution corresponding to a given parameter set (and any value of d), neither is it achieved at the lowest possible level of inequality (the constant equilibrium solution).

Parameter combinations which yield the latter behaviour were sought using a random search technique. An example solution (Set 2 in 6.4) which clearly exhibits a unique maximum mean equilibrium wealth is shown in Figure 6.10. It is also apparent that the maximum mean equilibrium wealth is not achieved at the highest level of inequality associated with equilibrium solutions for this parameter set and any value of d. This leads to the following insight.



(a) *Equilibrium mean wealth for different values of d*



(b) *Equilibrium normalised variance for different values of d*

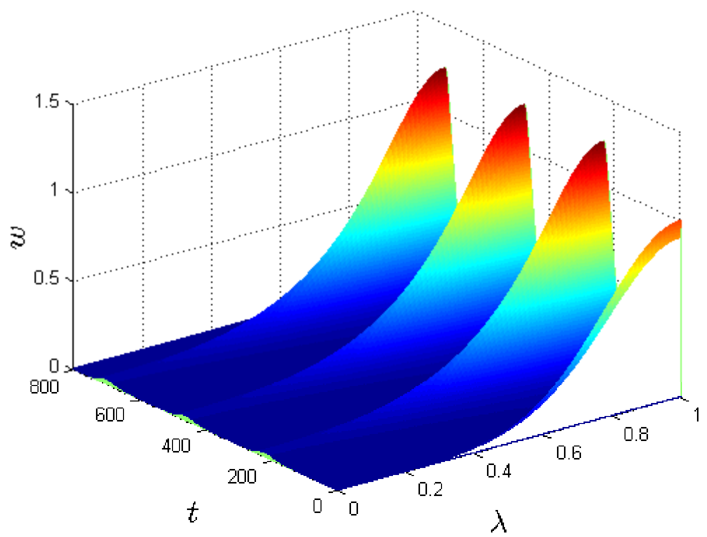
FIGURE 6.10: *The relationship between the redistribution rate, the mean wealth and the normalised variance at nonconstant equilibrium solutions to* (6.29)–(6.31) *for* $k = 4.85 \times 10^{-2}$ *and parameter set 2 in Table 6.4. The figures are based on numerical approximations of solutions to the model* (6.29)–(6.31) *obtained by applying the method* (3.29) *with step sizes* $\triangle t = 5 \times 10^{-3}$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$.

Observation 6.3 *In the context of increasing per capita wealth growth rates and decreased wealth growth prospects at a given fixed wealth level when either the mean wealth or the level of inequality increases, there may exist an optimal redistribution rate and associated level of inequality, which is not the highest level of possible stable inequality, for which the expected wealth of each entity in the economic system is maximised, and for which the associated equilibrium distribution is stable.*

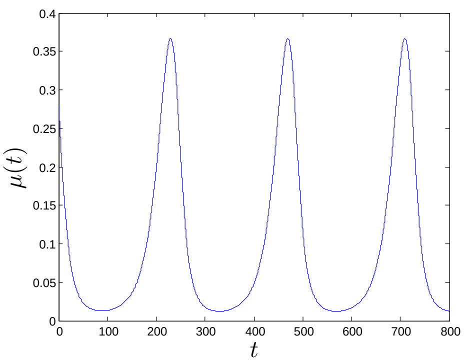
A conclusion drawn from the preceding section was that greater inequality at an equilibrium yields a greater, but less stable, mean wealth. This suggested that there exists a tradeoff between lowering inequality and seeking stability on the one hand, and maximizing the mean wealth on the other. The limiting case yielding the largest possible equilibrium wealth in the preceding section therefore forms the boundary of the basin of attraction for a given parameter set, and is hence unstable. The implication is that the maximum possible equilibrium wealth is unattainable if stability is required, and furthermore, a higher equilibrium wealth is attained only by accepting greater inequality.

Observation 6.3 suggests that this need not necessarily be the case. For the growth function considered in this section, an optimal redistribution rate may exist which maximises the mean wealth, at which the equilibrium is stable, and where the extent of inequality is less than the most unequal stable steady states.

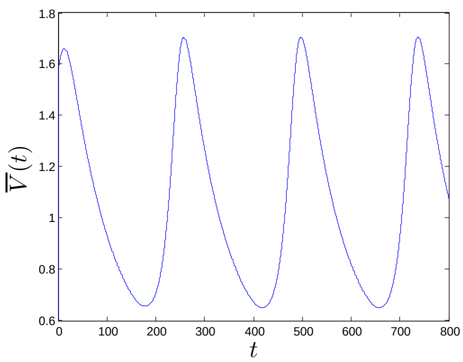
Attention is finally turned toward the oscillatory behaviour observed in Figure 6.8. There exist combinations of parameter sets and initial conditions for which growing oscillations over time are observed. An argument similar to that used to prove the existence of a critical redistribution rate in §6.1 suggests that there consequently exist cases which exhibit periodic oscillations over time. Figure 6.11 depicts such a structure over space and time, obtained using parameter set 3 of Table 6.4. This phenomenon leads to the following observation.



(a) *Wealth distribution over time*



(b) *Mean wealth over time*



(c) *Normalised variance over time*

FIGURE 6.11: *An example of periodically oscillating solution over time resulting from parameter set 3 in Table 6.4. The figures are based on numerical approximations of solutions to the model* (6.29)–(6.31) *obtained by applying the method* (3.29) *with step sizes* $\triangle t = 0.01$ *and* $\triangle\lambda = 1.25 \times 10^{-3}$, *and assuming the initial condition* $w_0(\lambda) = \exp(-10(\lambda - 1)^2)$.

Observation 6.4 *Evidently, linearly increasing per capita wealth growth rates (translating proportional to the level of inequality as well as to the total wealth) is sufficient to produce oscillating trends in the level of inequality as well as mean wealth in a wealth distribution over time, even in a society of identical agents without exogenously imposed growth influences.*

This suggests that cyclical (periodic) behaviour in socioeconomic systems need not necessarily be the result of explicit driving forces such as development, as originally conjectured by Kuznets [86]. As mentioned in §2.2, Kuznets' claims have been disputed, and the reverse of his postulate has even been claimed [9]. In light of the solution behavioural possibilities demonstrated for the model (6.29)–(6.31), either a 'u'-shaped curve or its inverse may be observed in both the extent of inequality as well as the total wealth, depending on the observation period.

6.4 Chapter summary

The long-time behaviour of the models of Chapter 4 were considered in this chapter. The per capita wealth growth-rate functions of §4.4 were incorporated into the models in order of increasing complexity. For the case of linear per capita wealth growth rates, the reliance of solution persistence on redistribution was demonstrated and a possible implication for the Robin Hood paradox was noted. The stability of the constant equilibrium solutions was analysed analytically and the existence of a nonconstant equilibrium solution was established in the case of the trickle-down model. The case of mean-dependent linear per capita wealth growth rates was considered next. It was demonstrated numerically that stable equilibrium solutions to the trickle-down model exist for a large class of model parameters. The relationship between the size of the basin of attraction and the redistribution rate was then demonstrated, as was the relationship between the mean equilibrium wealth and the redistribution rate.

Finally, the mean and inequality-dependent linear per capita wealth growth-rate function was considered and possible model solution behaviours were demonstrated numerically. The possibility of a finite optimal redistribution rate which maximises expected wealth was uncovered, as well as the possibility of oscillating solution behaviour over time (even in the absence of explicit oscillatory model parameter functions).

Part III

Conclusion

CHAPTER 7

Discussion

Contents

The results obtained in Chapters 4–6 are discussed in this chapter and related statements in the literature (in support or in contradiction of the findings of this thesis) are noted. First, the validity of the models underlying the arguments made in this thesis is considered and various limitations are noted in this respect. The result pertaining to solution persistence hinging on redistribution and its implications for the Robin Hood paradox is then considered. The fluctuating solution behaviour over time demonstrated in §6.3 and its implications for attempts at motivating trends in macroeconomics are noted next. The relation between oscillations and the extent of redistribution is also considered as a motivation for redistributive actions, as mentioned in §2.4. The possible existence of a unique, finite, optimal extent of redistribution is finally considered.

7.1 Model validation

The aim of the models adopted in this thesis was to explore theoretical possibilities, rather than capturing observed phenomena accurately. These models therefore cannot be validated in the traditional sense by comparing their results to actual observations in the hope of achieving a small enough discrepancy which may be deemed negligible.

For very simple cases, reasonably expected model solution behaviour could, however, be verified. For example, in the case of a constant per capita growth rate, wealth growth is independent of distributional characteristics, and it follows in this case that the total wealth represented by an aggregation of the model solution over space ought to grow exponentially over time regardless of the wealth distribution shape. This was verified analytically in §5.2.

The redistribution captured in the models of Chapter 4 is furthermore assumed to be conservative, that is, no wealth is lost or created in the act of redistribution itself. It is therefore to be expected that the change in the total wealth locally in time should not depend explicitly on redistribution. The redistribution parameter (the constant d) disappears for both models in the derivative of the total wealth (see Proposition 5.4), thus satisfying the aforementioned expectation.

Since redistribution by definition entails transfers that reduce the extent of both types of wealth inequality considered in this thesis, it is expected that there should always exist a redistribution rate which ensures that the extent of inequality decreases. This was confirmed in Propositions 5.1 and 5.2 for the two types of redistribution considered. Remarkably, a redistribution rate that ensures decreasing relative inequality can be found which involves only the extremal values of the wealth distribution (see Proposition 5.2). Although elegant and appealing in its simplicity, the significance of this sufficient condition should not be overstated since it depends strongly on the specific wealth growth-rate functional shapes assumed in the models.

A wealth distribution's shape is preserved when it is multiplied by a constant and for this reason relative inequality is invariant under constant per capita growth. It therefore follows that performing redistribution in respect of an unequal wealth distribution within the context of constant per capita growth should always result in decreased relative inequality. This expectation was confirmed in Proposition 5.1.

The expectations that follow naturally for the simplest possible type of wealth growth rate are therefore fulfilled by the models of this thesis. Furthermore, the model solution behaviours exhibited in the presence of more interesting types of growth, although at times initially surprising, do not defy logical explanation. The means of validation available in this type of investigation therefore indicate that the models analysed in this thesis are logically coherent and valid in the sense of producing solution behaviours deemed plausible under the associated assumptions.

Since the models considered in this thesis employ per capita growth-rate functions which are special cases of (4.16), no distinction has to be made when discussing the linear per capita growth models of §6.1 or the mean-dependent per capita growth models of §6.2 in this chapter. The possibilities in terms of their solution behaviour are included when considering (4.16) and choosing the appropriate parameters equal to zero.

It is important to note that since the aim in this thesis was to show what is theoretically possible in certain simple contexts, no predictive or explanatory statements related to reality can be made in respect of the models of Chapter 4. Solution behavioural possibilities can be uncovered and possible theoretical justifications for certain policies can be suggested, but no more.

7.2 The Robin Hood paradox

In §2.3, the Robin Hood paradox was reviewed, as were two irreconcilable motivations for its existence ([82, 83] *versus* [70, 79]). This emphasizes the importance of viewing prescriptive theories aimed at explaining economic phenomena in a sceptical light.

It was demonstrated in §6.1 that the persistence of wealth may hinge on the rate of redistribution (for both types of redistribution considered) and that there therefore exist cases where a lower bound on the possible extent of redistribution exists if the (global) persistence of wealth is required. If such a bound exists, there may exist scenarios in which the Robin Hood paradox necessarily follows. If the growth prospects in a particular economy are dire, such that only a small portion of the total entities in the economic system are expected to be able to increase their wealth positions, then the aforementioned bound on redistribution is larger than in the case where more entities have access to more promising growth prospects. If it is then taken into account that high levels of economic inequality have been associated with poor growth prospects [8, 51, 108, 112, 112], then it may be concluded that less redistribution ought to transpire in more unequal societies, given that (at least) the collective persistence of wealth is sought.

No claim is made that the Robin Hood paradox has been resolved or explained. It has simply been shown that the occurrence of such a contentious phenomenon may follow logically from very simple and weak assumptions. Its existence therefore does not imply the existence of driving forces (motivations other than the assumptions made in the previous argument), as is often assumed in *so called* explanations for manifestations of the Robin Hood paradox.

7.3 Economic fluctuation

It was demonstrated in §6.3 that fluctuating wealth and inequality levels over time can occur in the trickle-down model. That this behaviour is possible in a system of homogeneous entities with no time-dependent growth influences may at first be surprising, but such behaviour is reminiscent of those exhibited by autonomous mechanical systems, such as pendulums. In view of this analogy, it is no longer surprising that either vanishing, constant or growing fluctuations may be produced in the trickle-down model, since such behaviours are easily produced for a simple pendulum under the appropriate assumptions (weak damping, no damping and negative damping, respectively).

This very simple model is therefore already capable of producing rich enough solution behaviour to produce phenomena described by both Kuznets-type curves [86] and their inverses [9]. The possibility of such fluctuations in this simple model illustrates that fluctuating trends in macroeconomic metrics need not be the result of exogenously dictated or explicitly time-dependent growth processes.

The relationship between redistribution and fluctuating wealth is mentioned in passing by Hochman and Rodgers [74] in one of the first treatises of optimal redistribution, but was not considered in the ensuing analytical arguments. They suggested that greater redistribution will lead to greater income stability and therefore hedge persons against the uncertainty associated with future income fluctuations. Interdependent personal utility, rather than the aforementioned intuition, was used to justify redistributive actions theoretically.

In Figure 6.8 (interpreted together with Figure 6.9 (a)) it is apparent that in the presence of vanishing oscillations, greater redistribution is indeed related to faster vanishing oscillations within the context of the trickle-down model, which is consistent with the suggestion in [74]. It was also demonstrated in §6.2 that greater redistribution may be associated with greater stability (a larger basin of attraction of equilibrium solutions). The trickle-down model is therefore an example of an analytical model formulation which gives rise to a second theoretical justification for redistributive actions, entirely independent of the considerations of traditional utility theory.

7.4 Equality and optimal redistribution

It is commonly assumed that utility functions are concave over consumption or income [46], as mentioned in §2.5. In the simplest possible case, where it is assumed that all individuals share the same utility function, this leads to the conclusion that total equality of wealth maximises social welfare. More redistribution would then always be preferable to less redistribution. It was noted in §2.4 that the introduction of more heterogeneity, such as different levels of individual productivity, for example, can lead to an unequal distribution of wealth which maximises welfare, and which may therefore be called optimal [101]. If the assumptions on the shapes of utility functions and the combined social welfare function are, however, rejected, then there exist infinitely many Pareto optimal wealth distributions. The question of how much to redistribute therefore seems difficult to answer in the absence of assumptions pertaining to utility.

All equilibrium solutions to the trickle-down model are Pareto optimal allocations of wealth, since any increase (decrease) in the amount of redistribution adversely affects the wealth position of entities at the top (bottom) of the wealth distribution. If the mean wealth is considered as well, then a unique maximum in terms of mean wealth can be found, at a finite amount of redistribution, as demonstrated in §6.3 and particularly in Figure 6.10 (a). Some redistribution may therefore yield greater total wealth than either no or excessively severe redistribution. This is an example of a theoretical justification for limited redistributive intervention which is Pareto optimal in terms of allocation, and uniquely optimal in terms of maximising the total wealth. A single 'best' redistributive strategy is therefore obtained in this case without having to assume anything about individual utility. This approach is, however, not necessarily superior or preferable to arguments which rely on utility theory. In fact, a Rawlsian argument may be adopted to criticise this solution for rendering a subset of the population worse off in pursuit of increased total wealth, instead of minimising the disadvantages of the least well off. That the latter aim will arise naturally from the original position relies on the natural risk-averse nature of most humans [145], and is therefore in a sense reliant on the notion of utility as well. A perfectly rational (risk-neutral) decision maker who wishes to maximise his expected wealth will prefer the suggested optimal strategy. Regardless of these considerations, the primary aim was to demonstrate yet another way in which redistributive actions may be justified theoretically without appealing to utility theory (or, in this case, considerations related to stability).

The discussions in this and the preceding section refer only to the trickle-down model. The nonexistence of nonconstant equilibrium solutions to the linear redistribution model is a result of insufficient adaptability in the redistribution scheme as the wealth distribution changes. This suggests that proportional contributions should not be fixed to an entity's position (wealth or income) and the total wealth only, but should also take the distributional shape into account, in order to have a stabilising effect on the temporal evolution of a wealth distribution.

7.5 Chapter summary

The implications of the arguments put forward in this thesis were clarified in this chapter within the context of the relevant literature. First the validity and limitations of the models underlying the arguments of this thesis were discussed. It was then suggested that the Robin Hood paradox may, in fact, follow necessarily from fairly simple assumptions. The presence of oscillatory behaviour in the trickle-down model was also addressed. First, it was noted that such behaviour can result in this very simple model without exogenous influences. Secondly, the relationship between redistribution, fluctuations in wealth, and the stability of equilibria was suggested as an analytical formulation of a possible theoretical justification for redistributive actions. The existence of a unique finite extent of redistribution which maximises total wealth was then considered, and another possible justification for redistributive actions was suggested. A possible implication of the nonexistence of nonconstant equilibrium solutions to the linear redistribution model was finally noted.

CHAPTER 8

Summary and appraisal

Contents

This chapter comprises three sections. In the first, the contents of this thesis are summarised, while the second contains an appraisal of and a critique on the contributions made in the thesis. The final section is dedicated to the documentation of ideas for possible follow-up work related to the thesis contributions which may be pursued in the future.

8.1 Thesis summary

The discourse on economic inequality was reviewed very briefly in Chapter 2 from the point of view of social justice, in fulfilment of Objective I(a) of §1.3. The recurring theme of limited egalitarianism in terms of resource ownership was highlighted in this review. Empirically observed historic trends in economic inequality and certain resulting postulates, such as Kuznets' famous curve, were then described. The notion of redistribution of wealth was briefly reviewed by discussing various means of redistribution, the desirability of these various means, and the relationship between wealth redistribution and economic inequality. The Robin Hood paradox, which pertains to the relationship between the extent of economic inequality in a society and the extent of redistribution present therein, as well as differing explanations for the existence of this paradox, was also noted. The notion of optimal redistribution and theoretical justifications of redistributive acts were then considered. It was noted that existing theoretical justifications of redistributive acts rely heavily on utility theory. The measurement of economic inequality was discussed next and two important classes of indices for measuring inequality were reviewed. Existing models of wealth distribution were finally reviewed very briefly and the aim of this thesis was elucidated within the context of the literature review presented in the second chapter.

Certain mathematical prerequisites required to follow the investigation conducted in this thesis were reviewed in Chapter 3, in fulfilment of Objective I(b). Various mathematical definitions and theorems used in the arguments of this thesis were reviewed in the first section. A brief introduction to the field of differential equations was then given, after which a standard classification of second-order linear PDEs was reviewed. The focus of the discussion next shifted to reaction-diffusion systems in particular. A method for analysing the stability of equilibrium solutions to certain initial-boundary value problems involving PDEs was then reviewed. Certain numerical solution techniques for the approximation of solutions to initial-boundary value problems involving PDEs were then described briefly, after which emphasis was placed on the class of finite difference methods, since this was the numerical solution approximation scheme employed in later chapters.

Two novel mathematical models of wealth distribution were put forward in Chapter 4, in fulfilment of Objective II. First, the underlying model assumptions were noted and motivated, after which the mathematical model derivations were carried out. The per capita wealth growth-rate functions to be considered later in this thesis were then introduced. The existence and uniqueness of solutions to the aforementioned models were established next and it was demonstrated that these solutions, as well as their first derivatives in space, remain nonnegative for all time. Two metrics for characterising the extent of wealth equality present in model solutions were then defined, in fulfilment of Objective III, and their equivalence with traditional measures of wealth distribution was demonstrated. This was followed by a brief discussion on the interpretation and use of these metrics. The limitations and intended use of the models introduced were finally discussed critically and clarified in a formal model apology.

In Chapter 5, sufficient conditions for bounded wealth inequality were established, in fulfilment of Objective IV. The models derived in Chapter 4 were validated by analysing their solution behaviour in the context of a constant per capita wealth growth rate. An expression for the change in total wealth over time was derived for all per capita growth-rate function types considered later in the thesis. A necessary and sufficient condition for decreasing relative inequality was established next and an easily verifiable sufficient condition for decreasing relative inequality was derived from this necessary and sufficient condition, which depends only on the extremal values of the current wealth distribution. A necessary and sufficient condition for decreasing absolute inequality was also established and the differences between this condition and that for relative inequality were noted. Numerical examples illustrating some of the analytic results of the chapter were finally presented.

The long-time behaviour of solutions to the models of Chapter 4 were considered in the sixth chapter, in fulfilment of Objective V. The per capita wealth growth-rate functions of §4.4 were considered in order of increasing complexity. In the case of linear per capita wealth growth rates, a relationship between solution persistence and redistribution was demonstrated and a possible implication for the Robin Hood paradox was noted. The stability of the constant equilibrium model solutions was established analytically and the existence of a nonconstant equilibrium solution was established in the case of the trickle-down model. The case of mean-dependent linear per capita wealth growth rates was considered next. It was demonstrated numerically that stable equilibrium solutions to the trickle-down model exist for a large class of model parameters. A relationship between the size of the basin of attraction and the redistribution rate was then demonstrated, as well as a relationship between the mean equilibrium wealth and the redistribution rate. Finally, mean and inequality-dependent linear per capita wealth growth-rate functions were considered and possible solution behaviours were demonstrated numerically for these model incarnations. The possibility of a finite redistribution rate which maximises expected wealth was uncovered, as well as the possibility of oscillating solution behaviour over time.

The implications of the findings of the preceding chapters were finally clarified in the seventh chapter, in fulfilment of Objective VI. First, the validity and limitations of the underlying models employed to arrive at the arguments put forward in this thesis were discussed. It was then suggested that the Robin Hood paradox may necessarily follow from fairly simple assumptions. The presence of oscillatory solution behaviour in the case of the trickle-down model was also addressed. First, it was noted that such behaviour can result from this very simple model without exogenous or time-dependent influences. Secondly, a relationship between redistribution, fluctuations in wealth, and the stability of equilibria was suggested as an analytical formulation of a possible theoretical justification for redistributive actions. The existence of a unique, finite extent of redistribution which maximises total wealth was considered next, and this was suggested as another possible justification for redistributive actions. Finally, a possible implication of the nonexistence of nonconstant equilibrium solutions to the linear redistribution model was noted.

8.2 Appraisal and critique of thesis contributions

Two models were put forward in this thesis which represent examples of very simple analytic contexts capable of bringing forth mathematical descriptions of certain basic economic phenomena. Specifically, the notion of increasing (relative) inequality over time was pursued as an inherent characteristic of wealth growth in free market societies, as was the notion of wealth redistribution as some transfer of wealth from more wealthy entities to less wealthy entities.

Consideration of increasing per capita wealth growth-rate functions (increasing over wealth), is one way of capturing increasing relative inequality over time, the very simplest special case of this being linearly increasing per capita wealth growth-rate functions, which were investigated. Other than the notion related to increasing inequality over time, the inherent relative value of wealth (and hence competition to obtain more than the average) as well as the idea that increases in inequality reduce growth prospects, where captured in the formulation of wealth growth.

Two examples of redistribution dynamics were investigated. One example mimics diffusion-like effects, called trickle-down redistribution, and the other represents a conservative linear-tax transfer scheme.

It was demonstrated that increases in inequality can always be limited in the context of these models by means of sufficient redistribution. It was also demonstrated that there exist cases where the persistence of the total wealth hinges on the rate of redistribution, and that from this occurrence, the Robin Hood paradox may follow. Furthermore, it was illustrated that oscillating behaviour over time in wealth and inequality can occur even in the very simple contexts of the models presented in this thesis, in the absence of explicitly time-dependent underlying processes.

It was shown that redistributive actions can be motivated or justified in various ways other than in terms of the notion of utility, and that these arguments can be formulated analytically. Whether or not an argument can be formulated analytically is an excellent test of its coherence. Furthermore, it was illustrated that there can exist a single, finite redistribution rate which maximises the expected wealth of an entity (above the wealth level associated with an equal distribution), at a stable equilibrium.

The models in this thesis are far removed from reality. The scope of what can be inferred from their solutions is therefore limited in definite ways. The functional forms assumed in this thesis for the growth of wealth and its redistribution only represent examples of choices which capture certain phenomena. The solution behaviours of these models therefore only represent possibilities rather than predictions. The majority of results presented are merely demonstrations of why one should be highly sceptical of, and very careful in, the formulation of narratives that closely fit empirical observations in macroeconomics and complicated systems of interacting agents in general.

The aim in this thesis was not to be overly critical of existing explanations for economic phenomena related to wealth inequality, such as Kuznets' hypothesis. Instead, the aim was rather to attempt to convince the reader that these explanations should always be viewed merely as possibilities, or contributing factors, rather than definitive expositions of the causes and effects of economic processes.

8.3 Future work

In this final section, four possible avenues for future research following on the work documented in this thesis are suggested, in fulfilment of Objective VII.

Suggestion 8.1 *An analytic treatment of equilibrium solution stability classification.*

The primary aim in the analysis of the long-time behaviour of solutions to the models of Chapter 4 was to demonstrate possible solution behaviours, rather than pursue complicated analytic results related to equilibrium solution stability. For this reason, analytic results related to equilibrium solution stability were provided only in the very simplest case, namely for constant equilibrium solutions.

The stability properties of the nonconstant equilibrium solution to the trickle-down model may be pursued in future. It was demonstrated numerically that in the case of a simple linear per capita wealth growth-rate function, this solution is unstable. For the two other per capita wealth growth-rate functions, it was illustrated that there exist parameter sets for which nonconstant equlibrium solutions to the trickle-down model are indeed stable. In the latter case, in addition to establishing these stability properties analytically, results related to the size of the basin of attraction (exactly where bifurcations occur that separate solutions which tend to the same equilibrium solution from other solutions) may be pursued.

Suggestion 8.2 *A model solution classification for the case of* (4.16) *with* $c \neq 0$ *and* $k \neq 0$.

In the case of mean and inequality-dependent per capita wealth growth rates, certain interesting solution behaviours of the trickle-down model were uncovered during numerical experiments. An exhaustive classification of all possible behaviours was, however, not performed. Such a classification may be pursued upon application of scaling techniques aimed at reducing the number of model parameters to a minimum and performing a systematic search based on numerical experiments, possibly with the aid of an automated stability classification technique.

Suggestion 8.3 *An extension of the models to allow for multiple interacting society cases.*

The very simplest case of a single, closed economy was considered in this thesis. By allowing the wealth function w to be a vector function and incorporating assumptions related to inter-society competition or collaboration, the models of Chapter 4 may be extended to an n society case. In this context, the entire system's stability becomes a question of interest.

Such multiple society models may furthermore be solved in respect of specific network structures, where only adjacent societies are allowed to interact, in which case the effect of the specific network structure on the system's stability becomes a question of interest.

Suggestion 8.4 *A generalisation of the basic model assumptions of this thesis.*

In order to render the investigation of this thesis tractable, examples of functional forms that capture certain phenomena (such as increasing wealth inequality over time, or the relative value of wealth) were assumed. A more difficult, but perhaps more enlightening, approach might be to investigate the implications of the basic assumptions only. For example, assuming that the per capita wealth growth-rate function is any increasing function on wealth. Such an investigation may be preceded by first investigating more types of per capita wealth growth rate functional forms which are increasing over wealth in order to gain insight into commonalities between solutions to models employing increasing per capita growth rates.

The necessary and sufficient conditions for decreasing inequality derived in Chapter 5 applied to both types of redistribution considered in Chapter 4. Interesting solution behaviours at later stages in the investigation were, however, limited to only one of the models. This leads to the question of what can be said in general about any redistribution scheme based on the principle of transfers, in the context of increasing inequality as an inherent feature of wealth growth. Such an investigation may be preceded by first investigating other abstractions of redistribution than the two types considered in this thesis in order to gain insight into the essential characteristics of redistribution.

References

[1] Abramowitz M & Stegun IA, 1964, *Handbook of mathematical functions — With formulas, graphs, and mathematical tables*, United States Department of Commerce, Washington (DC).

[2] Achdou Y, Buera FJ, Lasry JM, Lions PL & Moll B, 2014, *Partial differential equation models in macroeconomics*, Philosophical Transactions of the Royal Society A, **372(2028)**, (no page numbers).

[3] Aghion P & Bolton P, 1997, *A theory of trickle-down growth and development*, The Review of Economic Studies, **64(2)**, pp. 151–172.

[4] Aghion P, Caroli E & Garcia-Penalosa C, 1999, *Inequality and economic growth: The perspective of the new growth theories*, Journal of Economic Literature, **37(4)**, pp. 1615–1660.

[5] Ahluwalia MS, 1976, *Inequality, poverty and development*, Journal of Development Economics, **3(4)**, pp. 307–342.

[6] Aiyagari SR, 1994, *Uninsured idiosyncratic risk and aggregate saving*, The Quarterly Journal of Economics, **109(3)**, pp. 659–684.

[7] Alesina A & Perotti R, 1996, *Income distribution, political instability, and investment*, European Economic Review, **40(6)**, pp. 1203–1228.

[8] Alesina A & Rodrik D, 1994, *Distributive politics and economic growth*, Quaterly Journal of Economics, **102(2)**, pp. 465–490.

[9] Anand S & Kanbur SR, 1993, *Inequality and development: A critique*, Journal of Development Economics, **41(1)**, pp. 19–43.

[10] Arnold DN, 2015, *Stability, consistency, and convergence of numerical discretizations*, pp. 1358–1364 in Engquist B (Ed), *Encyclopedia of applied computational mathematics*, Springer, Berlin.

[11] Atkinson AB, 1970, *On the measurement of inequality*, Journal of Economic Theory, **2(3)**, pp. 244–263.

[12] Atkinson AB & Harrison AJ, 1978, *Distribution of personal wealth in Britain*, Cambridge University Press, Cambridge.

[13] Avigad J & Donnelly K, 2004, *Formalizing O notation in Isabelle/HOL*, Proceedings of the International Joint Conference on Automated Reasoning, pp. 357–371.

[14] Ayres F, 1952, *Theory and problems of differential equations*, Schaum Publishing Company, New York (NY).

[15] Beaudry P, Blackorby C & Szalay D, 2009, *Taxes and employment subsidies in optimal redistribution programs*, The American Economic Review, **99(1)**, pp. 216–242.

[16] Bell SA, Henry JF & Wray LR, 2004, *A Chartalist critique of John Locke's theory of property, accumulation, and money: Or, is it moral to trade your nuts for gold?*, Review of Social Economy, **62(1)**, pp. 51–65.

[17] Bendix R, 1974, *Inequality and social structure: A comparison of Marx and Weber*, American Sociological Review, **39(2)**, pp. 149–161.

[18] Besley T & Coate S, 1991, *Public provision of private goods and the redistribution of income*, The American Economic Review, **81(4)**, pp. 979–984.

[19] Bewley T, 1986, *Stationary monetary equilibrium with a continuum of independently fluctuating consumers*, Contributions to Mathematical Economics in Honor of Gérard Debreu, **79**, pp. 79–102.

[20] Bhargava SC, 1989, *Generalized Lotka-Volterra equations and the mechanism of technological substitution*, Technological Forecasting and Social Change, **35(4)**, pp. 319–326.

[21] Blackorby C & Donaldson D, 1980, *A theoretical treatment of indices of absolute inequality*, International Economic Review, **21(1)**, pp. 107–136.

[22] Blackorby C & Donaldson D, 1978, *Measures of relative equality and their meaning in terms of social welfare*, Journal of Economic Theory, **18(1)**, pp. 59–80.

[23] Bleecker D & Csordas G, 1992, *Basic partial differential equations*, Van Nostrand Reinhold, New York (NY).

[24] Boadway R, Marchand M, Pestieau P & del Mar Racionero M, 2002, *Optimal redistribution with heterogeneous preferences for leisure*, Journal of Public Economic Theory, **4(4)**, pp. 475–498.

[25] Boyd JP, 2001, *Chebyshev and Fourier spectral methods*, 2nd Edition, Dover Publications Incorporated, Mineola (NY).

[26] Brams SJ, 1968, *Measuring the concentration of power in political systems*, American Political Science Review, **62(02)**, pp. 461–475.

[27] Brauer F, 1963, *Bounds for solutions of oridnary differential equations*, Proceedings of the American Mathematical Society, **14(1)**, pp. 36–43.

[28] Buchanan JM, 1959, *Positive economics, welfare economics, and political economy*, The Journal of Law and Economics, **2**, pp. 124–138.

[29] Buchanan JM, 1960, *Review: The theory of public finance*, Southern Economic Journal, **26(3)**, pp. 234–238.

[30] Burden R & Faires JD, 2011, *Numerical analysis*, 9th Edition, Prindle, Weber & Schmidt, Boston (MA).

[31] Butzer P & Weis R, 1977, *On the Lax equivalence theorem equipped with orders*, Journal of Approximation Theory, **19(3)**, pp. 239–252.

[32] Cantrell RS & Cosner C, 2004, *Spatial ecology via reaction-diffusion equations*, John Wiley & Sons, Chichester.

[33] Champernowne DG, 1952, *The graduation of income distributions*, Econometrica, **20(4)**, pp. 591–615.

[34] Chari V, Nicolini JP & Teles P, 2016, *More on the optimal taxation of capital*, [Online], [Cited February 2017], Available from `https://economicdynamics.org/meetpapers/2016/paper_1364.pdf`.

[35] Chatterjee A, Chakrabarti BK & Manna S, 2004, *Pareto law in a kinetic model of market with random saving propensity*, Physica A: Statistical Mechanics and its Applications, **335(1)**, pp. 155–163.

[36] Clarke GR, 1995, *More evidence on income distribution and growth*, Journal of Development Economics, **47(2)**, pp. 403–427.

[37] Clifford J, 2008, *Locating equality: From historical philosophical thought to modern legal norms*, Equal Rights Review, **1(1)**, pp. 11–22.

[38] Cowell FA, 2015, *Inequality and poverty measures*, Adler MD & Fleurbaey M (Eds), *The Oxford handbook of well-being and public policy*, Oxford University Press, Oxford.

[39] Cowell FA, 2000, *Measurement of inequality*, pp. 87–116 in Bourguignon F & Atkinson T (Eds), *Handbook of income distribution*, North-Holland, Oxford.

[40] Cowell FA & Kuga K, 1981, *Additivity and the entropy concept: An axiomatic approach to inequality measurement*, Journal of Economic Theory, **25(1)**, pp. 131–143.

[41] Creel HG, 1953, *Chinese thought: From Confucius to Mao Tse-tung*, The University of Chicago Press, Chicago (IL).

[42] Cropp R & Norbury J, 2015, *Population interactions in ecology: A rule-based approach to modeling ecosystems in a mass-conserving framework*, SIAM Review, **57(3)**, pp. 437–465.

[43] Dagum C, 1997, *A systematic approach to the generation of income distribution models*, Journal of Income Distribution, **6(1)**, pp. 105–126.

[44] Dagum C, 1977, *New model of personal income-distribution-specification and estimation*, Economie appliquée, **30(3)**, pp. 413–437.

[45] Dagum C, 2006, *Wealth distribution models: Analysis and applications*, Statistica, **66(3)**, pp. 235–268.

[46] Dalton H, 1920, *The measurement of the inequality of incomes*, The Economic Journal, **30(119)**, pp. 348–361.

[47] Daou J, 2016, *Stability theory*, Lecture Notes, University of Manchester, Manchester.

[48] Davidson JE, Hendry DF, Srba F & Yeo S, 1978, *Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom*, The Economic Journal, **88(352)**, pp. 661–692.

[49] Davies JB, Shorrocks A, Sandstrom S & Wolff EN, 2007, *The world distribution of household wealth*, Mapping Global Inequalities, Center for Global, International and Regional Studies, University of California, Santa Cruz (CA).

[50] Deacon R, 2003, *Dictatorship, democracy, and the provision of public goods*, Departmental Working Papers, Department of Economics, University of California, Santa Barbara (CA).

[51] Deininger K & Squire L, 1998, *New ways of looking at old issues: Inequality and growth*, Journal of Development Economics, **57(2)**, pp. 259–287.

[52] Düring B & Matthes D, 2010, *A mathematical theory for wealth distribution*, pp. 81–113 in Naldi G, Pareschi L & Toscani G (Eds), *Mathematical modeling of collective behavior in socio-economic and life sciences*, Springer, New York (NY).

[53] Epstein JM, 2008, *Why model?*, Journal of Artificial Societies and Social Simulation, **11(4)**, pp. 12–13.

[54] Eriksson K, Estep D & Johnson C, 2013, *Applied mathematics: Body and soul*, Springer Science & Business Media, New York (NY).

[55] Evans LC, 2010, *Partial differential equations*, 2nd Edition, American Mathematical Society, Providence (RI).

[56] Fair RC, 1979, *An analysis of the accuracy of four macroeconometric models*, The Journal of Political Economy, **87(4)**, pp. 701–718.

[57] Fisher RA, 1937, *The wave of advance of advantageous genes*, Annals of Eugenics, **7(4)**, pp. 355–369.

[58] Fisk PR, 1961, *The graduation of income distributions*, Econometrica, **29(2)**, pp. 171–185.

[59] Forbes KJ, 2000, *A reassessment of the relationship between inequality and growth*, American Economic Review, **90(4)**, pp. 869–887.

[60] Friedman M, 2009, *Capitalism and freedom*, University of Chicago Press, Chicago (IL).

[61] Furman J & Stiglitz JE, 1998, *Economic consequences of income inequality*, Proceedings of the Symposium Sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole (WY), pp. 221–263.

[62] Gallagher LJ, 1935, *Communism in Russia*, The Irish Monthly, **63(743)**, pp. 271–282.

[63] Gallegati M, Keen S, Lux T & Ormerod P, 2006, *Worrying trends in econophysics*, Physica A: Statistical Mechanics and its Applications, **370(1)**, pp. 1–6.

[64] Gastwirth JL, 1972, *The estimation of the Lorenz curve and Gini index*, The Review of Economics and Statistics, **54(3)**, pp. 306–316.

[65] Gemmell N & Morrissey O, 2005, *Distribution and poverty impacts of tax structure reform in developing countries: How little we know*, Development Policy Review, **23(2)**, pp. 131–144.

[66] Gilens M, 2012, *Affluence and influence: Economic inequality and political power in America*, Princeton University Press, Princeton (NJ).

[67] Greengard L & Rokhlin V, 1997, *A new version of the fast multipole method for the Laplace equation in three dimensions*, Acta Numerica, **6**, pp. 229–269.

[68] Guriev S & Rachinsky A, 2006, *The evolution of personal wealth in the former Soviet Union and Central and Eastern Europe*, World Institute for Development Economics Research (Research Paper No. 2006/120), United Nations University, Tokyo.

[69] Hale JK, 1971, *Functional differential equations*, Springer, New York City (NY).

[70] Hall PA & Soskice D, 2001, *Varieties of capitalism: The institutional foundations of comparative advantage*, Oxford University Press, Oxford.

[71] Hausman DM, 1995, *The impossibility of interpersonal utility comparisons*, Mind, **104 (415)**, pp. 473–490.

[72] Higgins M & Williamson JG, 2002, *Explaining inequality the world round: Cohort size, Kuznets curves, and openness*, Southest Asian Studies, **40(3)**, pp. 268–302.

[73] Hobbes T, 1982, *Leviathan*, Penguin Classics, London.

[74] Hochman HM & Rodgers JD, 1969, *Pareto optimal redistribution*, The American Economic Review, **59(4)**, pp. 542–557.

[75] Huggett M, 1993, *The risk-free rate in heterogeneous-agent incomplete-insurance economies*, Journal of Economic Dynamics and Control, **17(5)**, pp. 953–969.

[76] Humi M & Miller WB, 1992, *Boundary value problems and partial differential equations*, Prindle, Weber & Schmidt, Boston (MA).

[77] Irvin G, 2011, *Inequality and recession in Britain and the USA*, Development and Change, **42(1)**, pp. 154–182.

[78] Itô K, 1993, *Encyclopedic dictionary of mathematics*, MIT Press, Cambridge (MA).

[79] Iversen T & Soskice D, 2009, *Distribution and redistribution: The shadow of the nineteenth century*, World Politics, **61(03)**, pp. 438–486.

[80] Jha SK, 1996, *The Kuznets curve: A reassessment*, World Development, **24(4)**, pp. 773–780.

[81] Keefer P & Knack S, 2002, *Polarization, politics and property rights: Links between inequality and growth*, Public Choice, **111(1-2)**, pp. 127–154.

[82] Korpi W, 2006, *Power resources and employer-centered approaches in explanations of welfare states and varieties of capitalism: Protagonists, consenters, and antagonists*, World Politics, **58(02)**, pp. 167–206.

[83] Korpi W, 1985, *Power resources approach vs. action and conflict: On causal and intentional explanations in the study of power*, Sociological Theory, **3(2)**, pp. 31–45.

[84] Kuttler C, 2011, *Reaction-diffusion equations with applications*, Lecture Notes, Technische Universität München, München.

[85] Kutz JN, 2016, *Scientific computing*, Lecture Notes (Applied Mathematics 518), Univeristy of Washington, Washington (DC).

[86] Kuznets S, 1955, *Economic growth and income inequality*, The American Economic Review, **45(1)**, pp. 1–28.

[87] Leontief W, 1982, *Academic economics*, Science, **217(4555)**, pp. 104–107.

[88] Li H & Zou HF, 1998, *Income inequality is not harmful for growth: Theory and evidence*, Review of Development Economics, **2(3)**, pp. 318–334.

[89] Li N & Steiner J, 2001, *Finite difference schemes for n-dimensional reaction-diffusion equation*, Bulletin mathématique de la Société des Sciences Mathématiques de Roumanie, **44(3)**, pp. 221–236.

[90] Lindert PH, 2004, *Growing public: Social spending and economic growth since the eighteenth century*, Cambridge University Press, Cambridge.

[91] Lindert PH, 2000, *Three centuries of inequality in Britain and America*, pp. 167–216 in Bourguignon F & Atkinson T (Eds), *Handbook of income distribution*, North-Holland, Oxford.

[92] Loomis LH & Sternberg S, 1990, *Advanced calculus*, 2^{nd} Edition, Jones and Bartlett Publishers, Boston (MA).

[93] Ma WX & Fuchssteiner B, 1996, *Explicit and exact solutions to a Kolmogorov-Petrovskii-Piskunov equation*, International Journal of Non-Linear Mechanics, **31(3)**, pp. 329–338.

[94] MacCulloch R, 2005, *Income inequality and the taste for revolution*, Journal of Law and Economics, **48(1)**, pp. 93–123.

[95] Mahajan V, 1985, *Innovation diffusion*, John Wiley & Sons, Hoboken (NY).

[96] Malcai O, Biham O, Richmond P & Solomon S, 2002, *Theoretical analysis and simulations of the generalized Lotka-Volterra model*, Physical Review E, **66(3)**, (no page numbers).

[97] Mandelbrot B, 1960, *The Pareto-Levy law and the distribution of income*, International Economic Review, **1(2)**, pp. 79–106.

[98] Mason JC & Handscomb DC, 2002, *Chebyshev polynomials*, CRC Press, Boca Raton (FL).

[99] Meltzer AH & Richard SF, 1981, *A rational theory of the size of government*, The Journal of Political Economy, **89(5)**, pp. 914–927.

[100] Michau JB, 2014, *Optimal redistribution: A life-cycle perspective*, Journal of Public Economics, **111**, pp. 1–16.

[101] Mirrlees JA, 1971, *An exploration in the theory of optimum income taxation*, The Review of Economic Studies, **38(2)**, pp. 175–208.

[102] Moore EJ, 1985, *On system-theoretic methods and econometric modeling*, International Economic Review, **26(1)**, pp. 87–110.

[103] Morgan J, 1990, *Boundedness and decay results for reaction-diffusion systems*, SIAM Journal on Mathematical Analysis, **21(5)**, pp. 1172–1189.

[104] Mouhot C & Villani C, 2004, *Regularity theory for the spatially homogeneous Boltzmann equation with cut-off*, Archive for Rational Mechanics and Analysis, **173(2)**, pp. 169–212.

[105] Newell AC & Whitehead JA, 1969, *Finite bandwidth, finite amplitude convection*, Journal of Fluid Mechanics, **38(02)**, pp. 279–303.

[106] Okun AM, 2015, *Equality and efficiency: The big tradeoff*, Brookings Institution Press, Washington (DC).

[107] Olsen EO, 1969, *A normative theory of transfers*, Public Choice, **6(1)**, pp. 39–58.

[108] Panizza U, 2002, *Income inequality and economic growth: Evidence from American data*, Journal of Economic Growth, **7(1)**, pp. 25–41.

[109] Pareto V, 1964, *Cours d'économie politique*, Librairie Droz, Genéve.

[110] Pareto V, 1895, *La legge della domanda*, Giornale Degli Economisti, **10(6)**, pp. 59–68.

[111] Pareto V & Busino G, 1967, *Ecrits sur la courbe de la répartition de la richesse*, Librairie Droz, Genéve.

[112] Persson T & Tabellini G, 1994, *Is inequality harmful for growth?*, The American Economic Review, **84(3)**, pp. 600–621.

[113] Pestel R & Radermacher FJ, 2000, *Equity, wealth and growth: Why market fundamentalism makes countries poor*, [Online], [Cited April 2017], Available from `http://files.globalmarshallplan.org/equity_wealth_and_growth.pdf`.

[114] Pianegonda S, Iglesias JR, Abramson G & Vega J, 2003, *Wealth redistribution with conservative exchanges*, Physica A: Statistical Mechanics and its Applications, **322**, pp. 667–675.

[115] Pigou AC, 1912, *Wealth and welfare*, Macmillan and Company, London.

[116] Piketty T, 2015, *The economics of inequality*, Harvard University Press, Cambridge (MA).

[117] Piketty T, Goldhammer A & Ganser L, 2014, *Capital in the twenty-first century*, Harvard University Press, Cambridge (MA).

[118] Plato & Cornford FM, 1945, *The republic of Plato*, Oxford University Press, Oxford.

[119] Prasad N, 2008, *Policies for redistribution: The use of taxes and social transfers*, International Institute for Labour Studies (Discussion Paper No. 194), Genéve.

[120] Pye LW, 1991, *The state and the individual: An overview interpretation*, The China Quarterly, **127**, pp. 443–466.

[121] Ramsey FP, 1927, *A contribution to the theory of taxation*, The Economic Journal, **37(145)**, pp. 47–61.

[122] Rawls J, 2009, *A theory of justice*, Harvard University Press, Cambridge (MA).

[123] Rodriguez SB, Díaz-Giménez J, Quadrini V & Ríos-Rull JV, 2002, *Updated facts on the US distributions of earnings, income, and wealth*, Federal Reserve Bank of Minneapolis Quarterly Review, **26(3)**, pp. 2–32.

[124] Romer T, 1975, *Individual welfare, majority voting, and the properties of a linear income tax*, Journal of Public Economics, **4(2)**, pp. 163–185.

[125] Rothschild C & Scheuer F, 2016, *Optimal taxation with rent-seeking*, The Review of Economic Studies, **83(3)**, pp. 1225–1262.

[126] Rousseau JJ, 1984, *A discourse on inequality*, Penguin Books, London.

[127] Sagan H, 1974, *Advanced calculus*, Houghton Mifflin Company, Boston (MA).

[128] Samuelson PA, 1971, *Generalized predator-prey oscillations in ecological and economic equilibrium*, Proceedings of the National Academy of Sciences, **68(5)**, pp. 980–983.

[129] Saxonhouse AW, 1982, *Family, polity & unity: Aristotle on Socrates' community of wives*, Polity, **15(2)**, pp. 202–219.

[130] Sen A, 1992, *Inequality reexamined*, Clarendon Press, Oxford.

[131] Sheshinski E, 1972, *The optimal linear income-tax*, The Review of Economic Studies, **39(3)**, pp. 297–302.

[132] Simpson PLP, 2000, *The politics of Aristotle*, University of North Carolina Press, Chapel Hill (NC).

[133] Smith GD, 1985, *Numerical solution of partial differential equations: Finite difference methods*, Oxford University Press, Oxford.

[134] Solomon S & Richmond P, 2002, *Stable power laws in variable economies; Lotka-Volterra implies Pareto-Zipf*, The European Physical Journal B — Condensed Matter and Complex Systems, **27(2)**, pp. 257–261.

[135] Solomon S, 1999, *Generalized Lotka-Volterra (GLV) models and generic emergence of scaling laws in stock markets*, [Online], [Cited March 2017], Available from `https://arxiv.org/pdf/cond-mat/9901250.pdf`.

[136] Stewart J, 2012, *Calculus*, 7$^{\text{th}}$ Edition, Cengage Learning, Boston (MA).

[137] Stiglitz JE, 1969, *Distribution of income and wealth among individuals*, Econometrica, **37(3)**, pp. 382–397.

[138] Stiglitz JE, 1987, *Pareto efficient and optimal taxation and the new welfare economics*, pp. 991–1042 in Auerbach AJ, Chetty R, Feldstein M & Saez E (Eds), *Handbook of public economics*, Elsevier, Amsterdam.

[139] Süli E, 2012, *Finite element methods for partial differential equations*, Lecture Notes, Mathematical Institute, University of Oxford, Oxford.

[140] Tadmor E, 2012, *A review of numerical methods for nonlinear partial differential equations*, Bulletin of the American Mathematical Society, **49(4)**, pp. 507–554.

[141] Taleb NN, 2012, *Antifragile: How to live in a world we don't understand*, Random House, New York (NY).

[142] Tang S, Qin S & Weber R, 1993, *Numerical studies on 2-dimensional reaction-diffusion equations*, The Journal of the Australian Mathematical Society Series B: Applied Mathematics, **35(02)**, pp. 223–243.

[143] Taylor AE, 1958, *Introduction to functional analysis*, John Wiley & Sons, London.

[144] Thornton J, 2001, *The Kuznets inverted-U hypothesis: Panal data evidence from 96 countries*, Applied Economics Letters, **8(1)**, pp. 15–16.

[145] Tom SM, Fox CR, Trepel C & Poldrack RA, 2007, *The neural basis of loss aversion in decision-making under risk*, Science, **315(5811)**, pp. 515–518.

[146] Turing AM, 1952, *The chemical basis of morphogenesis*, Philosophical Transactions of the Royal Society of London B: Biological Sciences, **237(641)**, pp. 37–72.

[147] Vinci F, 1921, *Nuovi contributi allo studio della distribuzione dei redditi*, Giornale degli economisti e Rivista di statistica, **61(11)**, pp. 365–369.

[148] Watson GN, 1995, *A treatise on the theory of Bessel functions*, Cambridge University Press, Cambridge.

[149] Winckler EA, 1999, *Transition from communism in China: Institutional and comparative analyses*, Lynne Rienner Publishers, Boulder (CO).

[150] Yemtsov R, 2007, *Housing privatization and household wealth in transition*, World Institute for Development Economics Research (Research Paper No. 2007/02), United Nations University, Tokyo.

[151] Zill DG, 1986, *Differential equations with boundary-value problems*, 2$^{\text{nd}}$ Edition, PWS-KENT, Boston (MA).

[152] Zill D, Wright WS & Cullen MR, 2011, *Advanced engineering mathematics*, Jones & Bartlett Learning, Burlington (MA).

[153] Zoutman FT & Jacobs B, 2016, *Optimal redistribution and monitoring of labor supply*, Journal of Public Economics, **135**, pp. 15–31.